Analysis of protein-protein interaction by *in vivo* quantitative proteomics in *Caenorhabditis elegans*

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

M.Sc. Jiaxuan Chen

Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Richard Lucius

Gutachter/innen: 1. Prof. Dr. Richard Lucius

2. Prof. Dr. Matthias Selbach

3. Dr. Markus Landthaler

Tag der mündlichen Prüfung: 21.09.2015

for my parents

“No man is an island, entire of itself.”

John Donne

Table of Contents

Abstract

Systemic investigation of protein-protein interactions has been instrumental in elucidating the complex molecular mechanisms underlying animal development. In *C. elegans*, early embryogenesis is a dynamically regulated process and provides an attractive model system for mapping *in vivo* protein interactions. In order to accurately identify specific interactions in *C. elegans* embryos, a new quantitative approach was developed combining *in vivo* expressed GFP fusion proteins with label-free interaction proteomics. This strategy was applied to studying the interaction partners of eight bait proteins involved in essential biological processes during early embryogenesis. As a result, this study generated a pilot embryo *in vivo* interaction network composed of 559 interactions among 472 proteins. Importantly, this network captures not only well-characterized bindings but also new interactions of high functional relevance. Further utility of the network is demonstrated by combining it with RNAi perturbation to search for new regulators of P granule formation in early embryos. Consequently, a worm-specific protein GEI-12 was discovered as a novel interaction partner of the DYRK kinase MBK-2 and as an important regulator of P granule dynamics and germline maintenance. This leads to a hypothetical model in which the phosphorylation state of GEI-12 mediates P granule assembly and disassembly during early embryogenesis. In addition, GEI-12 also induces granule formation in mammalian cells and interacts with PP2A phosphatases, indicating that the fundamental biophysical properties required for ribonucleoprotein granule formation are conserved across species during evolution. In summary, *in vivo* interactome mapping is a powerful and versatile approach that not only unravels the functional organization of the proteome but also can reveal invaluable biological insights into animal development.

Zusammenfassung

Die systematische Untersuchung von Protein-Protein-Interaktionen hat wesentlich zum Verständnis der komplexen molekularen Mechanismen beigetragen, welche der Entwicklung von Tieren zu Grunde liegt. In *C. elegans* ist die frühe Embryogenese ein dynamisch regulierter Prozess und bietet ein attraktives Modellsystem, um Wechselwirkungen von Proteinen *in vivo* zu entschlüsseln. Zur präzisen Identifizierung von spezifischen Interaktionen im *C. elegans* Embryo wurde ein neuer quantitativer Ansatz entwickelt, welcher die Expression von Fusionsproteinen an grün fluoreszierendes Protein *in vivo* mit markierungsfreier Interaktionsproteomik kombiniert. Diese Strategie wurde angewandt, um die Interaktionspartner von acht Proteinen zu untersuchen, die in essentiellen biologischen Prozessen während der frühen Embryogenese involviert sind. Diese Studie liefert als Ergebnis ein erstes embryonales *in vivo* Interaktionsnetzwerk bestehend aus 559 Interaktionen zwischen 472 Proteinen. Dieses Netzwerk erfasst nicht nur bekannte Bindungen, sondern auch neue Interaktionen von hoher funktioneller Relevanz. Die Netzwerkinformationen wurden mit Experimenten auf Basis der Ribonukleinsäuren-Interferenz kombiniert um neue Regulatoren der sogenannten „P granules" (zelluläre Strukturen in der Keimbahn) ausfindig zu machen. Infolgedessen wurde das fadenwurmspezifische Protein GEI-12 als neuer Interaktionspartner der DYRK-Kinase MBK-2 und als wichtiger Regler für die Dynamik der „P granules" und für die Aufrechterhaltung der Keimbahn identifiziert. Dies führt zu einem hypothetischen Modell in welchem der Phosphorylierungszustand von GEI-12 den Auf- und Abbau der „P granules" während der frühen Embryogenese vermittelt. Darüber hinaus veranlasst GEI-12 auch die Entstehung von „P granules" in Säugetierzellen und bindet an PP2A-Phosphatasen, was darauf hindeutet, dass die grundlegenden biophysikalischen Eigenschaften die zur Entstehung der Ribonukleoprotein-Körperchen notwendig sind, im Laufe der Evolution zwischen Spezies konserviert geblieben sind. Zusammenfassend stellt die *in vivo* Interaktionskartierung ein leistungsstarkes und vielseitiges Werkzeug dar, welches nicht nur die funktionelle Organisation des Proteoms aufdeckt, sondern auch wertvolle biologische Einsichten in die tierische Entwicklungsbiologie liefert.

I Introduction

I.1. Protein-protein interaction

Protein is of paramount importance in executing messages stored in the genome. The proteome, which could be loosely defined as the entire collection of proteins expressed by an organism, is not a simple flow from the genome or transcriptome. The dynamic range of a cellular proteome spans seven orders of magnitude (Zubarev, 2013). Furthermore, the complexity of the proteome is considerably amplified by post-translational modifications (PTMs) of proteins, including phosphorylation, acetylation, methylation, glycosylation, ubiquitination, lipidation, proteolysis and many others. As a result, the tightly controlled yet intricate combination of different PTMs diversify protein functionalities, allowing fine-tuning of protein activity in response to various intra- and extra-cellular signals and biological needs.

Proteins rarely act on their own, and in fact, most proteins are believed to be associated with other proteins and biomolecules to exert a plethora of molecular functions. Although each cell is packed with a large number of proteins, the localization and coordination of different proteins are beautifully programmed and regulated not only in a cell-autonomous fashion but also according to the functional needs of a wider physiologic system. The interactions between different proteins are therefore critically important in completing desirable molecular tasks for essentially every biological process. Given the fundamental importance of protein-protein interactions, it is not surprising, when out of proper control, aberrant pathologic protein-protein interactions also form the basis that drives the progression of many diseases, including a broad spectrum of neurodegenerative disorders and cancers (Ivanov et al., 2013; Ross and Poirier, 2005).

To explore the functions of a protein, identifying its interaction partners gives a critical overview of the biophysical environment in which the protein exerts its functions. Such a picture allows for a mechanistic insight into the biological processes where these proteins are involved. At a systems level, comprehensive mapping of protein-protein interactions provides crucial details of the mechanical and functional organization of the proteome, which are instrumental in elucidating complex molecular phenomena.

The goal of the present thesis is to develop a new quantitative proteomic strategy to study protein-protein interactions *in vivo* during early embryogenesis of *Caenorhabditis elegans*, with a biological sub-focus on P granule dynamics. Therefore, in the next sections, current methodology to investigate protein-protein interactions will be introduced. This is followed by an overview of P granule regulation during *C. elegans* early development and a brief summary of current knowledge on regulatory mechanisms of ribonucleoprotein (RNP) granule assembly/disassembly in other systems.

I.1.1. Detection of protein-protein interactions

From a physical chemistry point of view, protein-protein interaction is mediated primarily through non-covalent intermolecular forces, including electrostatic interactions, hydrogen bonding, van der Waals forces, cation-π interactions and hydrophobic effect (Crowley and Golovin, 2005; Leckband, 2000). Despite being relatively weak individually, multiple interactions can occur at the binding interface and generate strong association between proteins. There're various ways to classify protein-protein interactions into different categories, based on their composition, binding affinity and stability (Ozbabacan et al., 2011). According to the lifetime of their association, protein-protein interactions are generally described as stable or transient. Despite being conceptually different, they are also sometimes interchangeably perceived as strong or weak interactions, which in fact refer to their binding affinities. Well-characterized protein complexes are usually formed by stable interactions that maintain their integrity throughout lengthy biochemical purification, for example, the histone octameric core complex (Eickbush and Moudrianakis, 1978) and the core components of ribosomes and proteasomes (Fromont-Racine et al., 2003; Gu and Enenkel, 2014). However, many dynamic interactions are transient and weakly associated, and they are heavily involved in diverse cellular processes, such as signal transduction and protein synthesis (Gallie, 2002; Ozbabacan et al., 2011). Identifying these interactions poses a technically more challenging task.

Over the last decades, new techniques for detecting protein-protein interactions have continually been developed. Apart from computational prediction and microscopic colocalization assays, biochemical and biophysical methods are widely used for qualitative and quantitative assessment of protein bindings.

Targeted approaches which require prior knowledge of the proteins to be tested, such as co-immunuoprecipitation followed by immunoblotting, far-Western blotting, surface plasmon resonance and isothermal titration calorimetry, are generally limited to testing a small number of binding candidates, whereas high-throughput techniques, such as protein microarray, phage display and yeast-two hybrid, allow fast detection in a multiplexed format and can be expanded to a genome-wide scale.

Currently, many of the systemic protein-protein interaction networks of pathogens and model organisms have been generated using yeast-two hybrid system (Giot et al., 2003; LaCount et al., 2005; Li et al., 2004; McCraith et al., 2000; Rain et al., 2001; Rual et al., 2005; Stelzl et al., 2005; Uetz et al., 2006; Uetz et al., 2000). While these maps of binary interactions have contributed significantly to systems-level understanding of the proteomes, some inherent drawbacks of the technique are also noticeable (Parrish et al., 2006). Particularly, such a heterologous system requires two foreign proteins to interact in the yeast cell nucleus. This is one of the reasons for the significant number of interactions that are non-biologically relevant in the yeast-two hybrid data, i.e. the false positive problem (Parrish et al., 2006). Another perhaps under-estimated limitation is that the binary nature of yeast-two hybrid renders it not optimal for detecting cooperative bindings in protein complexes (Whitty, 2008). Therefore, an approach that allows proteins to interact in their native cellular environment and unbiased detection would be an improvement to the methodology.

I.1.2. Affinity purification and mass spectrometry

In contrast to targeted methods, a fully exploratory alternative to discover protein interaction partners is affinity purification coupled to mass spectrometry (Gingras et al., 2007). Typically, antibodies are used to capture the bait protein together with its binding partners from cell lysates, followed by mass spectrometry analysis (Fig. I.1). In an unbiased manner, this powerful approach allows identifying multiple protein complexes in one single assay. Yet, one of the biggest challenges is to discriminate specific interactions from non-specific contaminants that are bound to the same affinity matrix. This problem becomes even more pronounced with the ever-increasing sensitivity and speed of modern mass spectrometry instruments.

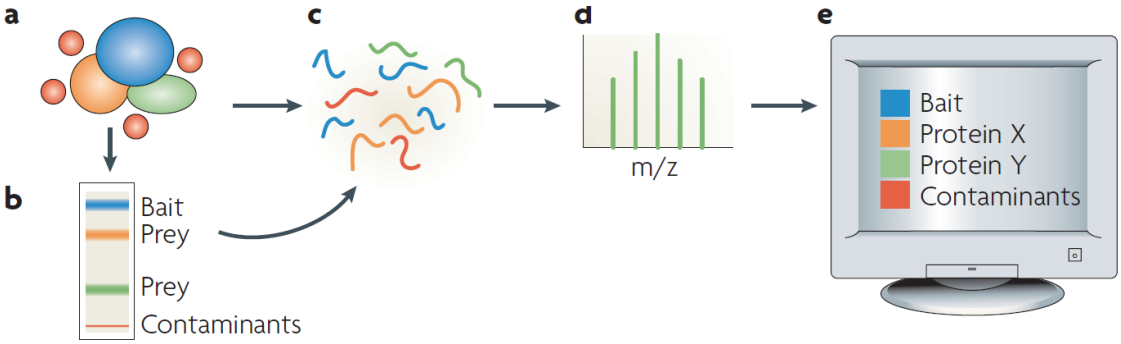

Fig. I.1 | General workflow of a typical affinity purification and mass spectrometry experiment for protein complex identification. Illustration reproduced from Gingras et al. (2007).

I.1.2.1 Tandem affinity purification

To circumvent the issue of non-specific bindings, one of the early proposed solutions is the tandem affinity purification (TAP) strategy (Rigaut et al., 1999). This strategy was successfully used to map the interactome of budding yeast *Saccharomyces cerevisiae* (Gavin et al., 2002). As depicted in Fig. I.2, the TAP tagging approach employs multiple purification steps. Ideally, the final eluate should be clean from background binders. However, this stringent purification strategy can only partially eliminate background contaminants while increasing the risks of losing interactions that are biologically relevant but low in binding affinity (Vermeulen et al., 2008). Therefore, a newer approach was expected to address these problems.

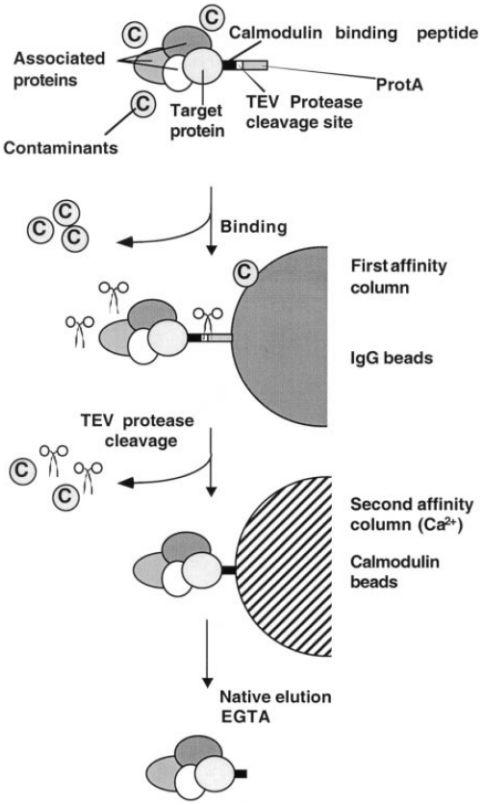

Fig. I.2 | Overview of the tandem affinity purification (TAP) tagging approach. The TAP approach employs a tandem affinity tag that allows protein complex purification in two steps in order to minimize bindings from background contaminants. Scheme taken from Puig et al. (2001).

I.1.2.2 Affinity purification and quantitative proteomics

In order to preserve weak interactions while still being able to distinguish specific binders from background contaminants, single-step low stringency affinity purification combined with quantitative proteomics is currently the method of choice (Paul et al., 2011; Vermeulen et al., 2008). As shown in Fig. I.3, in this approach, cells are metabolically labeled with different stable isotopes, typically using SILAC (stable isotope labeling by amino acids in cell culture) (Ong et al., 2002). Following affinity purification, samples are combined to minimize technical variations during further processing steps. The co-purified proteins are not only identified but also quantified by mass spectrometry. Importantly, the mass difference introduced by different isotope labels makes it possible to differentiate proteins originating from different cell populations in a complex mixture.

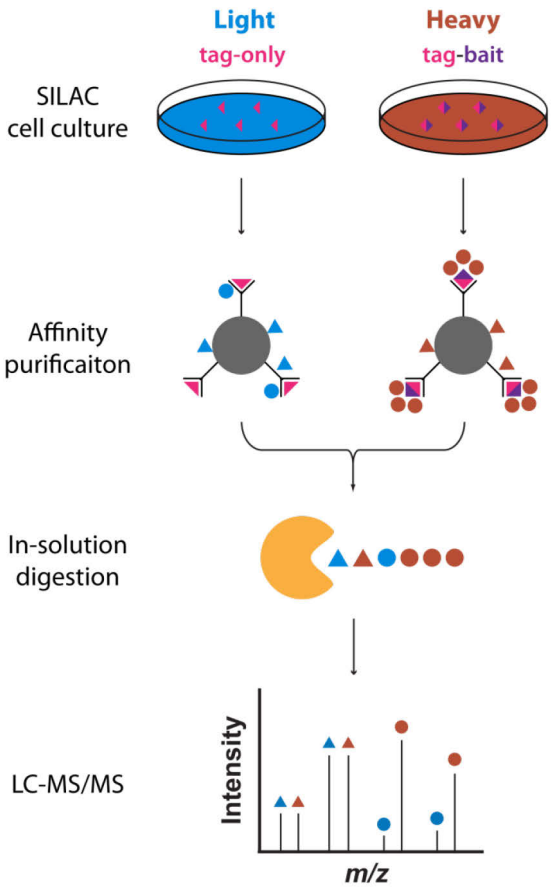

Fig. I.3 | Overview of affinity purification with SILAC-based quantitative proteomics using an exogenous affinity tag approach. Affinity-purified proteins of different SILAC states are combined and measured in the same mass spectrometry assay. Specific interaction partners can be identified by quantitative analysis of enriched bindings to the bait protein. LC-MS/MS, liquid chromatography-tandem mass spectrometry.

Furthermore, the quantification step assists in the search for proteins specifically enriched in the experiment pull-down in comparison to the negative control. With this information, specific interaction partners can be assigned with high confidence while background binders can be identified based on their similar abundances in both the experiment and control pull-downs. Of note, although typically not as accurate as labeling approaches, label-free quantification has also been successfully employed for quantitative analysis of affinity pull-downs (Hubner et al., 2010; Hubner and Mann, 2011; Rinner et al., 2007; Smits et al., 2013). While such

a low stringency approach helps to preserve weak interactions, it would also retain more background binders. Therefore, a carefully designed negative control is crucial for a successful experiment.

I.1.3. Negative control strategies

In order to accurately identify background contaminants among affinity-purified proteins, a properly designed control experiment is a vital prerequisite. Depending on the set-up of the experiment, various strategies for the control may be applied (Fig. I.4). Generally, the blocked empty bead approach would only work if the antibody used for the experiment pull-down has minimal cross-reactivity to undesired antigens, which must be carefully tested. Another commonly used approach is antibody isotype control, i.e., to use an antibody of the same isotype as that in the experiment pull-down but against an unrelated epitope. The assumption for using this strategy is that the non-specific bindings of the isotype control antibody will be the same as the experiment antibody except for their different target antigens. However, the two different antibodies could behave very differently from expectation, which could lead to wrong assignment of binding specificity.

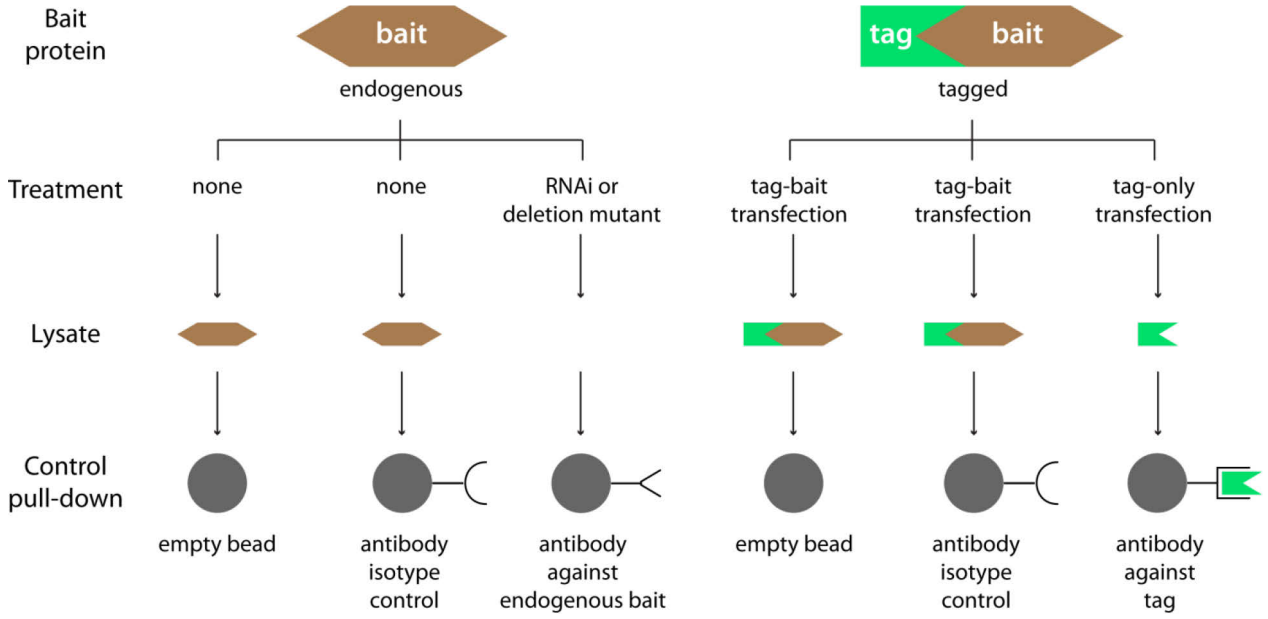

Fig. I.4 | General strategies for the negative control in quantitative affinity purification experiments.

In order to use the same antibody for both pull-downs, such methods as QUICK (quantitative immunoprecipitation combined with knockdown), which depletes the endogenous expression of the bait

protein in the control, can capture the same profile of non-specific binders as in the experiment pull-down (Fig. I.4) (Selbach and Mann, 2006). This control approach, however, is not applicable when the bait protein corresponds to an essential gene or when its depletion would lead to a substantial change of the proteome. Currently, the tag-only control strategy combined with transient transfection is widely used for cell-based quantitative affinity purification experiments (Paul et al., 2011). This approach makes use of well-characterized antibodies against affinity tags for both the experiment and control pull-downs. Also, transient transfection offers the advantage of minimizing background changes between the control and experiment cell populations, as compared to potential clonal differences among stably integrated cell lines.

I.1.4. *In vivo* protein-protein interactions

Protein-protein interaction screens using yeast-two hybrid or cell lines have generated interaction networks for organisms from simple life forms to humans (Boxem et al., 2008; Ewing et al., 2007; Giot et al., 2003; Li et al., 2004; Malovannaya et al., 2011; Rual et al., 2005; Simonis et al., 2009; Stelzl et al., 2005; Uetz et al., 2000). In spite of the scale of these interaction maps, *in vitro* studies can only capture a certain fraction of the complex interlinked system between proteins of a living organism. For instance, protein expression is regulated according to the functional needs of different tissues. In a developing multi-cellular organism, protein expression is also dynamically regulated. Thus, some interacting proteins identified through yeast-two hybrid or cell lines might in fact never be co-expressed in the same cellular compartment within the same tissue. Such interactions are not physiologically relevant *in vivo*. What are equally important are the post-translational modifications (PTMs) which can also mediate protein-protein interactions (Li et al., 2013a). In an *in vitro* system, these context-dependent modifications may not be correctly reflected as *in vivo*, thus leaving some biologically important interactions undetected. Consequently, the limitations of *in vitro* approaches necessitate further investigation of the complex interplay of proteins *in vivo*.

Affinity purification and mass spectrometry has emerged as a powerful approach also for studying protein-protein interactions *in vivo*. Most of the previous studies, including a recent report on *Drosophila* interactome using more than 200 knock-in lines, were performed in a non-quantitative manner, which renders it very difficult to correctly assign genuine interactions (Angrand et al., 2006; Bartoi et al., 2010;

Cheeseman et al., 2004; Lowe et al., 2014; Rees et al., 2011). A few *in vivo* studies that employed quantitative interaction proteomics however have a limited scale to single bait proteins (Andlauer et al., 2014; Hanack et al., 2015). In order to effectively map protein-protein interactions *in vivo*, substantial practical and technical challenges remain to be tackled. For instance, antibodies that can efficiently capture endogenous protein complexes are not always readily available, especially for non-mammalian model organisms. Similarly, in order to take advantage of affinity tags, it would require generating a large collection of knock-in transgenic animals. In parallel, due to multiple cell types that are present in a tissue sample, proper quantification and control strategies are particularly critical for the correct assignment of specific interactions.

Given the limitations of current methodology, a new approach is needed for generating *in vivo* protein-protein interaction networks. The large resource of GFP fusion transgenic lines of *C. elegans* (both for visualization and affinity purification) plus accurate label-free quantification (Cox et al., 2014) provide a unique opportunity to develop such a method that is able to accurately identify the endogenous binding partners of the bait protein in a living organism. Particularly, the embryos of *C. elegans* appear to be an attractive model system for investigating *in vivo* protein-protein interactions. In the next section, an overview of *C. elegans* embryogenesis will be presented with a focus on the regulation of P granule dynamics during early development.

I.2. P granule biology in a nutshell

I.2.1. Germline development during early embryogenesis

The microscopic nematode roundworm of *Caenorhabditis elegans* is an extensively researched model organism for life science. Thanks to its translucent anatomy, highly reproducible pattern of embryonic cell divisions and a wide range of genetic toolkits, *C. elegans* has been heavily used in studies of animal development. Among the many branches of developmental biology, one of the most intriguing is the oocyte-to-embryo transition. Upon fertilization, the fusion of two differentiated gametes (oocyte and sperm,

originating from primordial germline stem cells) triggers complex reprogramming to become yet again a totipotent zygote which actively divides and eventually develops into a complete organism. The onset of early embryogenesis kick-starts a series of rapid and tightly regulated molecular events which drive among others the segregation of germline lineage (Robertson and Lin, 2013).

As shown in Fig. I.5, the entry of a sperm into an oocyte lays down the cue for the posterior pole of the embryonic axis (Goldstein and Hird, 1996) and initiates egg activation (Marcello and Singson, 2010). The newly fertilized egg which was initially arrested at meiotic prophase then completes the remaining cycle of meosis I (Yamamoto et al., 2006). After undergoing two rounds of meiotic divisions, the two pronuclei fuse together and the resulting nucleus then enters the embryonic mitotic cell cycle (Robertson and Lin, 2013). The first mitosis is the first asymmetric cell division which splits the zygote (P0) into the somatic blastomere (anterior) and the germline precursor (P1, posterior). As embryonic cell division progresses, the germline blastomere is further segregated and remains as a single precursor cell (P1 -> P2 -> P3 -> P4) until around 100-cell stage when it divides into Z2 and Z3 primordial germ cells (Sulston et al., 1983). At late first larval stage, these two Z2 and Z3 progenitor cells start proliferating and eventually grow into the two gonad arms of germ cells (Updike and Strome, 2010).

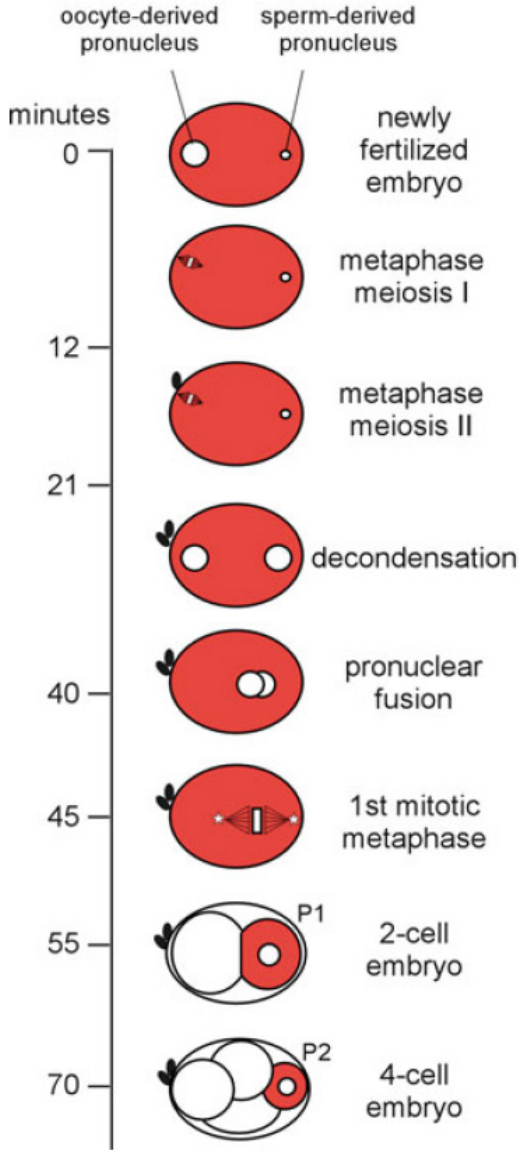

Fig. I.5 | Segregation of germline lineage during *C. elegans* early embryogenesis. Germline blastomeres are shown in red. Reproduced from Robertson and Lin (2013).

I.2.2. P granule segregation during oocyte-to-embryo transition

Germ plasm is the region of maternally derived cytoplasm present in the oocyte which during embryogenesis is inherited by only some cells of the embryo and will direct these cells to the germ cell fate (Wylie, 2000). Germ plasm has been discovered in multiple animal species. In *C. elegans*, suspected germ plasm was first described as electron-light cytoplasmic areas close to the nucleus of the germline precursor cell in early-stage embryos (Krieg et al., 1978). These microscopically visible granular structures were later termed P granules after being discovered by fluorescence microscopy using monoclonal antibodies (Strome and Wood, 1982, 1983) and under electron microscope (Wolf et al., 1983). Due to its segregation with the germline, P granules have traditionally been used as a marker for tracking germ cell lineage throughout *C. elegans* development (Fig. I.6).

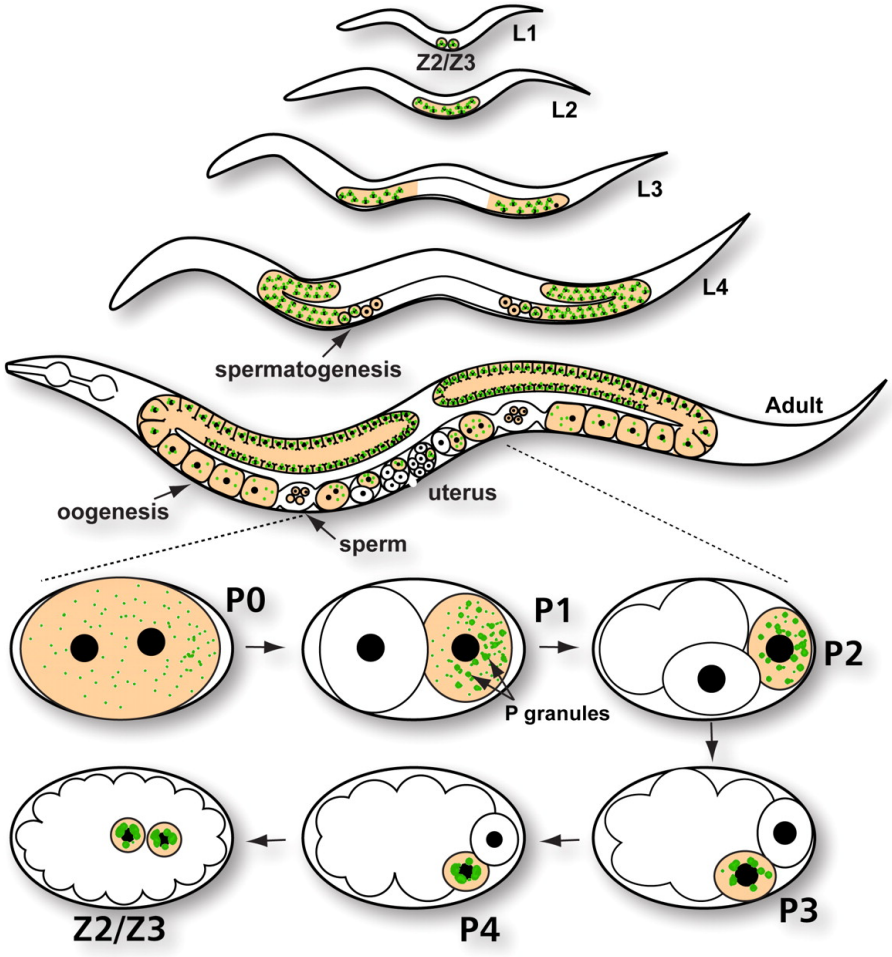

Fig. I.6 | Diagram of P granule segregation during *C. elegans* development. P granules, depicted in green, consistently segregate with the germline lineage. Reproduced from Updike and Strome (2010).

In almost all the gonadal germ cells of *C. elegans*, P granules show a characteristic peri-nuclear localization pattern where they are associated with nuclear pores (Pitt et al., 2000). Exceptions are in the maturing oocytes where P granules dissociate from the nuclear envelope and become evenly distributed in the cytoplasm as small aggregates (Fig. I.6). P granules are also present in developing spermatocytes but become undetectable in mature sperms (Updike and Strome, 2010). In the one-cell embryo, during pronuclear migration, P granules start to grow in size and localize towards the posterior pole. By the time of pronuclear fusion, the majority of P granules are already portioned into the posterior half of the embryo (Hird et al., 1996). Segregation of P granules with the germline lineage continues as embryonic development progresses.

I.2.3. P granule composition and function

Over 40 proteins have been identified to be P granule components (Table I.1), mainly based on microscopic colocalization experiments (Sengupta et al., 2013; Updike and Strome, 2010). While some of these proteins, such as GLH-1, GLH-2, PGL-1 and PGL-3, are constitutive components of P granules, others are only transiently associated with P granules during the life cycle of a worm (Updike and Strome, 2010). Importantly, all P granule components are known to be involved in RNA metabolism and/or possess RNA-binding domains in their amino acid sequences (Updike and Strome, 2010). Indeed, gonadal P granules have been found to contain RNA (Pitt et al., 2000). Embryonic P granules also harbor maternally loaded mRNA which is protected from degradation (Seydoux and Fire, 1994).

Based on its segregation with germ cells and the features of individual components, P granule is thought to be involved in germline specification/differentiation and play important roles in post-transcriptional regulation in the germline (Pitt et al., 2000; Updike and Strome, 2010). However, the function of P granule remains largely unclear. A previous study reported that in the *pptr-1* mutant P granules (as shown by GLH and PGL proteins) failed to assemble and therefore were not present in the germline blastomeres during early embryogenesis, but that the progeny were still fertile (Gallo et al., 2010). This suggests that P granules are not required for primordial germ cell determination. However, the segregation of other asymmetrically distributed proteins, such as MEX-5 and PIE-1, remained unaltered (Gallo et al., 2010). In another study,

Table I.1 | Known *C. elegans* P granule protein components §

Protein	Description
CAR-1	Cytokinesis, apoptosis, and RNA-binding 1 TRAL/Lsm14
CCF-1	CCR4/NOT deadenylase complex
CDE-1	Uracil nucleotidyltransferase
CGH-1	Dhh1/DDX6 DEAD-box helicase
CSR-1	Argonaute required for endo-siRNA
DCAP-1	mRNA decapping enzyme
DCAP-2	mRNA decapping enzyme
DEPS-1	Novel defective P granules and sterile
DRH-3	Dicer related DEAD-box helicase
EGO-1	RNA-directed RNA polymerase
GLD-1	RNA-binding KH domain
GLD-2	Poly(A) polymerase
GLD-3	RNA-binding KH domain
GLD-4	Poly(A) polymerase
GLH-1	Vasa DEAD-box helicase
GLH-2	Vasa DEAD-box helicase
GLH-3	Vasa DEAD-box helicase
GLH-4	Vasa DEAD-box helicase
GLS-1	Novel GLD-3/4 interacting protein
IFE-1	eIF4E mRNA cap-binding
IFET-1	Translational repressor
LAF-1	DDX3 DEAD-box helicase
MEG-1	Novel maternal effect germ cell defective
MEG-2	Novel maternal effect germ cell defective
MEX-1	CCCH-type zinc-finger protein
MEX-3	RNA-binding KH domain
OMA-1	CCCH-type zinc-finger protein
OMA-2	CCCH-type zinc-finger protein
PAB-1	Poly(A)-binding protein 1
PATR-1	Pat1-decapping cofactor
PGL-1	Novel RGG domain
PGL-2	Novel PGL-1 related
PGL-3	Novel RGG domain
PIE-1	CCCH-type zinc-finger protein
POS-1	CCCH-type zinc-finger protein
PRG-1	Argonaute required for piRNA synthesis
Sm proteins	Splicing factors
SPN-2	eIF4E-binding protein
SPN-4	RNP-type RNA-binding domain
TIA-1	TIA-1 RNP-type RNA-binding domain
VBH-1	Vasa Belle-like DEAD-box helicase
WAGO-1	Argonaute required for endo-siRNA

§ Adapted and summarized from Updike and Strome (2010) and Sengupta et al. (2013). siRNA, small interfering RNA; piRNA, Piwi-interacting RNA.

when multiple P granule components were simultaneously depleted, germ cells in the adult gonads lost their totipotency and initiated somatic reprogramming, which indicates P granules are essential for maintaining the germ cell fate (Updike et al., 2014). These authors proposed a model that P granules selectively degrade or compromise the translation of certain soma-promoting transcripts, thereby inhibiting the somatic fate and preserving the germ cell identity (Updike et al., 2014).

I.2.4. Regulation of P granule dynamics during early embryogenesis

Many genes that influence the formation of P granules have been identified through mutagenesis and RNAi studies (Updike and Strome, 2009). During early embryonic development, P granules undergo formation and dissolution in a highly dynamic yet tightly controlled manner (Fig. I.7). In the newly fertilized egg which remains unpolarized, P granules are evenly distributed as small aggregates. Signals from the sperm-derived centrosome trigger polarization and break the symmetry of the one-cell embryo. During polarization, posterior PAR proteins (PAR-1, PAR-2, LGL-1) accumulate at the posterior cortex; whereas, the anterior PAR proteins (PAR-3, PAR-6, PKC-3) retract from the posterior, consequently leading to two domains of PAR proteins in the polarized embryo (Hoege and Hyman, 2013). In parallel, the RNA-binding protein MEX-5 and its paralog MEX-6, which were uniformly diffuse in the cytosol before symmetry breaking, start building a gradient (high in the anterior, low in the posterior) and eventually become localized only to the anterior half (Schubert et al., 2000). As a consequence, P granules begin the localization towards the posterior half (around the male pronucleus) of the embryo and grow into bigger granules (Fig. I.7).

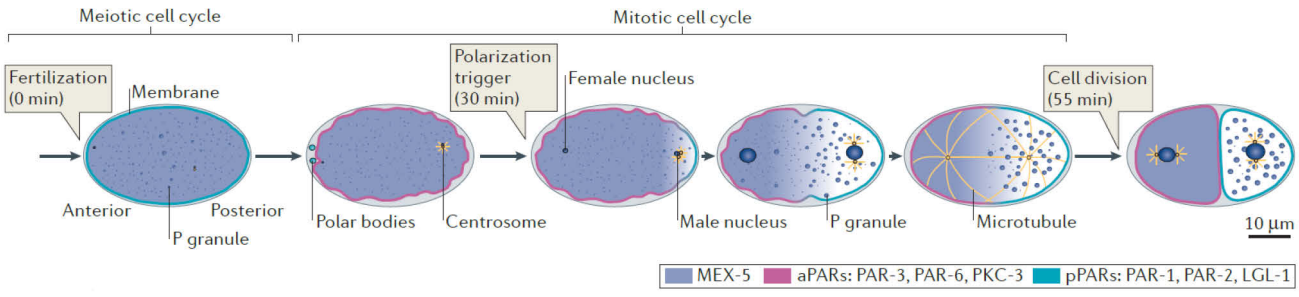

Fig. I.7 | Asymmetric partition of polarity proteins and P granules during *C. elegans* early embryogenesis. Adapted from Hoege and Hyman (2013).

P granules have been shown to possess liquid-like properties (Brangwynne et al., 2009). Contrary to the previous theory that posterior localization of P granules are caused by migration under cytoplasmic flow (Cheeks et al., 2004; Hird et al., 1996), Brangwynne and co-authors presented evidence that the flux of P granules into the posterior has a similar magnitude to that of the anterior flux, and therefore excluded the role of cytoplasmic flow in P granule partioning in the one-cell embryo (Brangwynne et al., 2009). Instead, these authors proposed that, after symmetry breaking, the condensation point of soluble P granule proteins is lowered in the posterior. As a result, this leads to posterior condensation of components released from dissolved anterior P granules, and this process is controlled via the gradient of MEX-5 regulated by asymmetric PAR-1 (Brangwynne et al., 2009). A slightly different model was proposed by Gallo et al. (2010), in which the regulatory mechanisms are independent for anterior dissolution during interphase (mediated by MEX-5) and posterior condensation during mitosis (promoted by PAR-1). They further discovered that posterior P granule assembly during mitosis was mediated via *pptr-1*, functionally downstream of PAR-1 (Gallo et al., 2010). The protein PPTR-1 is a regulatory subunit of protein phosphatase 2A (PP2A) family. Their finding that *pptr-1* is required for P granule assembly during mitosis suggests that dephosphorylation by PP2A may be important for P granule formation.

I.2.5. Other key players of anterior P granule disassembly?

In the one-cell embryo, the anteriorly polarized RNA-binding proteins MEX-5/6 have been repeatedly shown to mediate P granule disassembly. Yet, the mechanism how MEX-5 and MEX-6 control granule dissolution remains unknown (Nishi et al., 2008). During oocyte-to-embryo transition, asymmetric cell division is also in part regulated by selective partitions of maternal proteins (Fig. I.8) (Robertson and Lin, 2013). One of the key regulators that assist in this transition is MBK-2.

The kinase MBK-2 belongs to the dual-specificity tyrosine phosphorylation-regulated kinase (DYRK) family (Aranda et al., 2011). As shown in Fig. I.8, the activated MBK-2 is initially inhibited and anchored to the cell cortex of maturing oocytes by the catalytically inactive pseudophosphatase EGG proteins (Cheng et al., 2009). Upon fertilization, cortical MBK-2 is internalized into the cytoplasm of the embryo and appears as puncta. This step begins from the anaphase of meiosis I through meiosis II (Pellettieri et al., 2003). Regarding

molecular function, MBK-2 has been shown to phosphorylate a subclass of maternal proteins and consequently leads to their degradation, including OMA-1, OMA-2 and MEI-1, which are harmful for further development following the first embryonic cell division (Nishi and Lin, 2005; Stitzel et al., 2006). MBK-2 also phosphorylates MEX-5 and this modification primes MEX-5 for the subsequent phosphorylation by polo kinases PLK-1 and PLK-2, which is required for activation of MEX-5 *in vivo* (Nishi et al., 2008). In addition, segregation of some germline proteins, such as PIE-1, MEX-1 and POS-1, is also mediated by protein degradation in the somatic blastomeres, which is activated by MEX-5 and MEX-6 (DeRenzo et al., 2003).

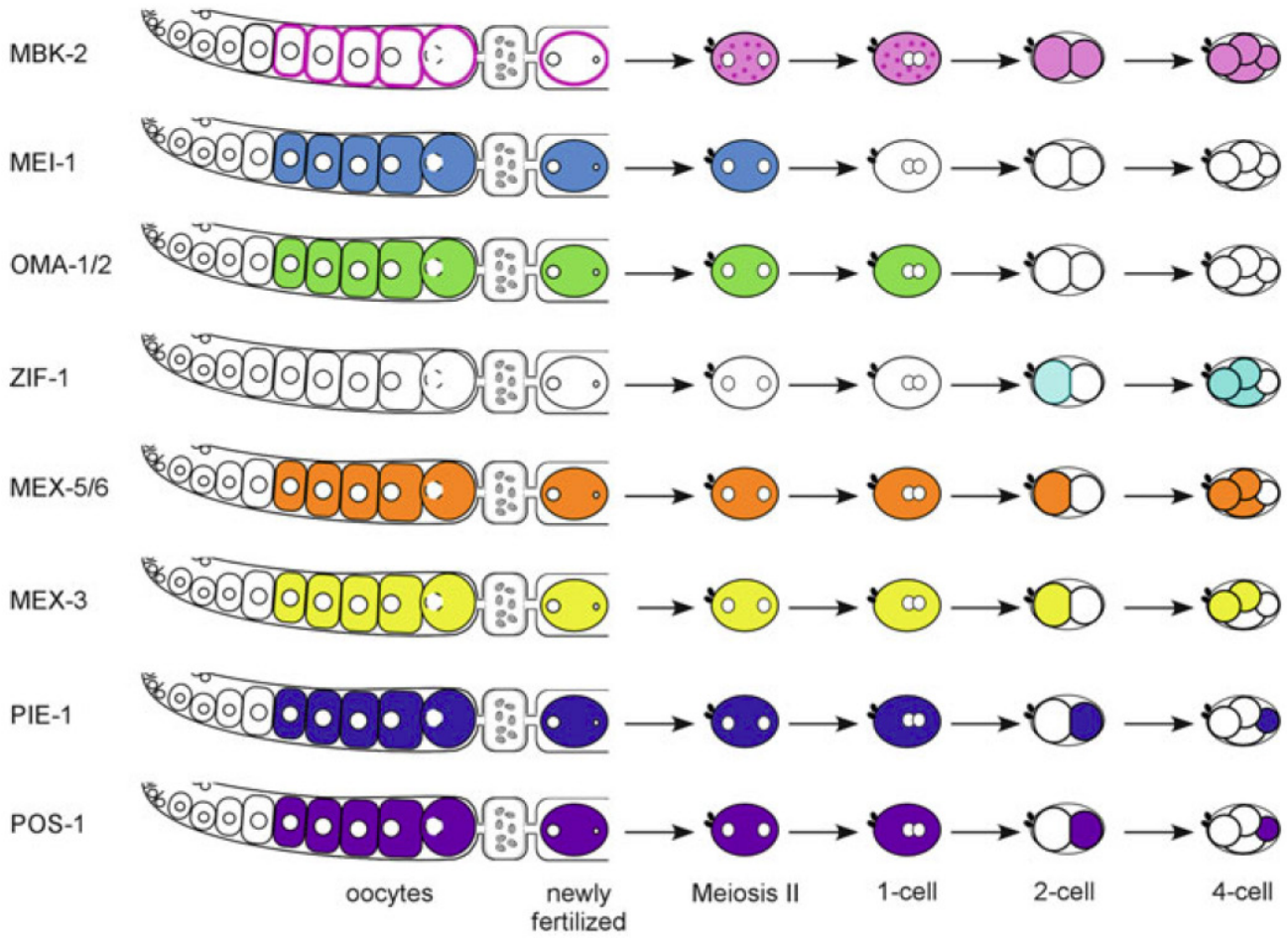

Fig. I.8 | Diagram showing the dynamic localization patterns of some maternal proteins during oocyte-to-embryo transition in *C. elegans*. Reproduced from Robertson and Lin (2013).

Given the critical importance of MBK-2 in regulating multiple processes during oocyte-to-embryo transition, MBK-2 might also directly or indirectly regulate anterior P granule dissolution. Indeed, depletion of *mbk-2* by RNAi leads to failure in asymmetric segregation of P granules (PGL-1) (Pellettieri et al., 2003). This phenotype was observed even in the mildly affected *mbk-2(RNAi)* embryos where microtubule-related

defects were not present (Pang et al., 2004), suggesting MBK-2 regulates P granule segregation via a microtubule-independent mechanism (Pellettieri et al., 2003). Finally, P granule proteins that remain in the somatic cells of the embryo have also been found to be degraded in the autophagosomes, mediated via cargo proteins SEPA-1 and EPG-2 (Tian et al., 2010; Zhang et al., 2009).

P granule is arguably the best-studied *in vivo* model system for ribonucleoprotein (RNP) granules. These granules are specific examples of a more general phenomenon of liquid-liquid demixing through phase transition. In the next section, a brief summary of known regulatory mechanisms of RNP granule formation in other systems will be presented.

I.3. Regulation of ribonulceoprotein granule assembly and disassembly in other systems

Apart from P granules, other types of non-membrane bound compartments exist ubiquitously in the nucleus and cytosol of eukaryotic cells (Hyman et al., 2014). Many of these membraneless compartments have been shown or predicted to resemble liquid droplets (Brangwynne et al., 2009; Brangwynne et al., 2011; Hyman et al., 2014). Examples include nucleoli, Cajal bodies, sites of DNA repair in the nucleus and various cytoplasmic RNP granules such as stress granules, processing bodies (P-bodies) and neuronal RNA granules (Hyman et al., 2014; Kiebler and Bassell, 2006; Weber and Brangwynne, 2012). RNP granules have been found to be involved in a wide variety of biological processes, such as RNA metabolism, translational regulation and stress response (Kiebler and Bassell, 2006; Weber and Brangwynne, 2012). However, the mechanisms that regulate the assembly and disassembly of these RNP granules remain not well understood.

I.3.1. Biophysical properties of proteins that promote liquid-liquid demixing

Through phase transition, the demixed liquid compartments allow rapid assembly of soluble components to build up high local concentrations and fast exchange with the surrounding pool of other molecules (Hyman et al., 2014). Likewise, when no longer required, these dynamic compartments can also be quickly dissembled (Hyman et al., 2014). Certain biophysical properties of proteins that promote liquid-liquid

demixing have been discovered (Weber and Brangwynne, 2012). Among them is multivalent binding (Li et al., 2012). Li and co-authors made use of a trimeric system in which the adaptor protein Nck binds to N-WASP, and, together with phosphorylated Nephrin, activates actin polymerization (Jones et al., 2006). By synthetically generating molecules that harbor multiple SH3 domains of Nck or multiple proline-rich motifs (PRM) of its ligand N-WASP, Li et al. demonstrated that these soluble molecules coalesced into liquid droplets in solution, depending on the number of the repeated domains. Furthermore, by co-expressing the two repeated domains (SH3 and PRM) in living cells, these authors also observed colocalized liquid-like puncta, suggesting multivalency could be a general mechanism governing liquid-phase separation (Li et al., 2012).

Another amino acid sequence feature that has been shown to drive liquid droplet formation is the low compositional complexity region (Han et al., 2012; Kato et al., 2012). In a cell-free system, Kato et al. showed that low complexity regions of some RNA-binding proteins could undergo concentration-dependent phase transition to form hydrogels. Furthermore, these hydrogels had an amyloid-like structure, but unlike typical amyloid aggregation their formation was dynamic and reversible, and they were able to retain the low complexity regions of other RNA-binding proteins as well as mRNA (Han et al., 2012; Kato et al., 2012).

I.3.2. Regulation of RNP granule disassembly by kinase activity

Primary amino acid sequence features are not the only mechanisms through which liquid-liquid demixing is controlled. Given the dynamic nature of RNP granules, control mechanisms that can rapidly regulate granule stability in response to cell signaling are expected. Indeed, the dissolution of stress granules, which are induced during various types of cellular stress conditions, has been found to require the dual-specificity kinase DYRK3 (Wippich et al., 2013). DYRK3 is a mammalian homolog of the *C. elegans* kinase MBK-2. Wippich and co-authors showed that the inactive form of DYRK3 localized to stress granules and prevented their dissolution. Upon activation of DYRK3, stress granules dissolved, leading to the release of the sequestered mTORC1 from the granules for downstream signal transduction. In addition, activated DYRK3 also directly phosphorylated the inhibitor protein PRAS40 to allow mTORC1 reactivation, thereby regulating the mTOR signaling pathway during stress conditions (Wippich et al., 2013).

I.4. Objectives of the thesis

Studying protein-protein interactions *in vivo* is arguably the most relevant approach to understanding the complex interplay of the proteome of an organism under a defined biological context. A promising method is to combine affinity purification using transgenic GFP fusion *C. elegans* strains with label-free quantitative proteomics as displayed in Fig. I.9. Therefore, the first aim of the present thesis is to develop a pipeline that allows efficient and quantitative analysis of *in vivo* protein interactions in *C. elegans*.

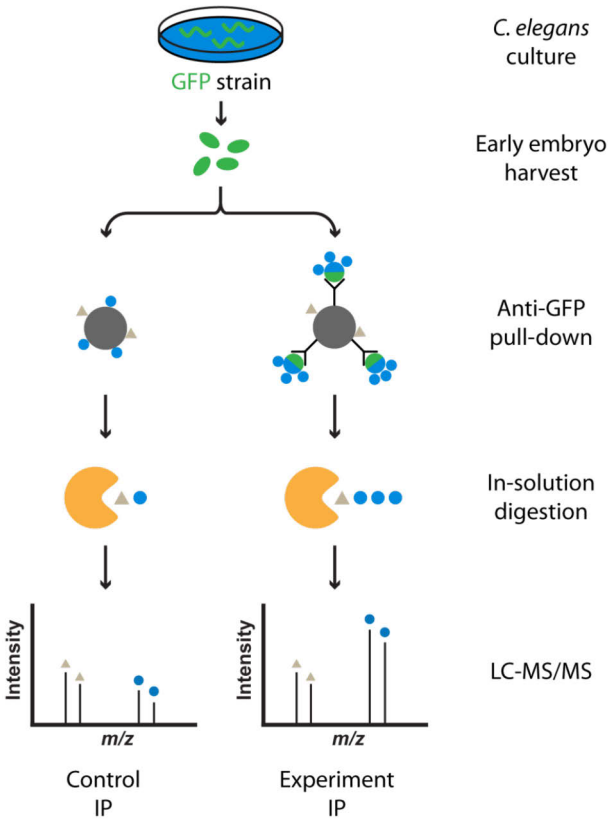

Fig. I.9 | Experimental strategy for affinity purification combined with label-free quantitative proteomics to identify specific interaction partners in *C. elegans* early embryos. LC-MS/MS, liquid chromatography-tandem mass spectrometry.

Early embryogenesis is one of the most intriguing branches of developmental biology. The smooth execution of this complex developmental program relies on intricate but tightly controlled regulatory mechanisms. In *C. elegans*, many of key players mediating early development have been identified through large-scale RNAi and mutagenesis screens (Fernandez et al., 2005; Fraser et al., 2000; Gonczy et al., 2000; Kemphues et al., 1988; Sonnichsen et al., 2005). To this end, a pilot map of *in vivo* interactions of essential proteins in early embryos would shed light on how these key player proteins coordinate the early stage of development. Such a network would also be a potentially important resource for a wider *C. elegans* research community.

Germline segregation during asymmetric cell divisions is a critical step in the embryonic development of *C. elegans*. Over the last decades, the association of P granules with the germline has been heavily investigated. Yet, many unanswered questions remain to be solved. By functionally following up the interaction partners of the bait proteins involved in P granule dynamics, novel regulatory mechanisms might

be discovered. These newly identified mechanisms might also help to understand the assembly and disassembly of RNP granules in general.

II Materials and Methods

II.1. Chemicals

Unless stated otherwise, all chemicals were purchased from the following suppliers: Carl Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), Life Technologies (Carlsbad, USA), Sigma-Aldrich (Steinheim, Germany).

II.2. Buffers and solutions

II.2.1. Microbiology culture

LB (lysogeny broth) medium/agar

Pre-mixed LB medium/agar powder was purchased from Sigma-Aldrich, dissolved in distilled water, sterilized by autoclaving.

Ampicillin

100 mg/ml stock solution in distilled water, sterilized by filtration, stored at -20 °C.

Kanamycin

50 mg/ml stock solution in distilled water, sterilized by filtration, stored at -20 °C.

Chloramphenicol

34 mg/ml stock solution in ethanol, sterilized by filtration, stored at -20 °C.

Tetracycline

50 mg/ml stock solution in ethanol, sterilized by filtration, stored at -20 °C.

II.2.2. DNA sample preparation and electrophoresis

6× DNA sample loading buffer

60 mM Tris-HCl pH 8.0, 60% [v/v] glycerol, 0.002% [w/v] bromophenol blue

TAE (Tris-acetate-EDTA) running buffer

40 mM Tris-HCl pH 8.5, 40 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA)

II.2.3. SDS-PAGE and Western blotting

5X SDS (sodium dodecyl sulfate) sample loading buffer

250 mM Tris-HCl pH7.4, 5mM EDTA, 10% [w/v] SDS, 50% [v/v] glycerol, 0.025% [w/v] bromophenol blue

MOPS (3-morpholinopropane-1-sulfonic acid) running buffer

40 ml 20x NuPAGE MOPS running buffer in 760 ml distilled water

Transfer buffer

25 mM Tris-HCl pH 8.5, 190 mM glycine, 20% [v/v] methanol in distilled water

TBST (Tris-buffered saline Tween-20) buffer

25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% [v/v] Tween-20 in distilled water

Blocking solution

5% [w/v] non-fat milk powder in TBST buffer.

Stripping buffer (Yeung and Stanley, 2009)

20 mM Tris-HCl pH7.4, 6 M guanidinium chloride, 0.2% [v/v] NP-40; 0.1 M β-mercaptoethanol freshly added.

II.2.4. LC-MS sample preparation

All solutions for LC-MS (liquid chromatography-mass spectrometry) applications were prepared in LC-MS grade (LiChrosolv) water or acetonitrile.

ABC (ammonium bicarbonate) buffer

50 mM ammonium bicarbonate (NH_4HCO_3) in water

Denaturation buffer

6 M urea, 2 M thiourea in 10 mM HEPES pH 8.0

Reduction buffer

10 mM dithiothreitol (DTT) in 50 mM ABC buffer

Alkylation buffer

55 mM iodoacetamide in 50 mM ABC buffer

Buffer A

5% [v/v] acetonitrile, 0.1% [v/v] formic acid in water

Buffer A*

5% [v/v] acetonitrile, 3% [v/v] trifluoroacetic acid in water

Buffer B

80% [v/v] acetonitrile, 0.1% [v/v] formic acid in water

II.2.5. *C. elegans* culture and maintenance

M9 buffer

22 mM KH_2PO_4, 42 mM Na_2HPO_4, 86 mM NaCl, 1mM $MgSO_4$ in distilled water, sterilized by autoclaving.

Phosphate buffer

1 M stock solution (0.8 M KH_2PO_4 , 0.2 M K_2HPO_4, pH 6.0), sterilized by autoclaving.

Nematode growth media (NGM)

Pre-mixed NGM powder was purchased from US Biologicals (Salem, USA), dissolved in distilled water, sterilized by autoclaving; when cooled down to 58 °C, phosphate buffer (25 mM), CaCl2 (1 mM) and $MgSO_4$ (1 mM) were added (to final concentration); aliquots were stored at 4 °C.

Bleach solution

NaOCl (available chlorine 1~1.5%), 0.5 M NaOH, 150 mM NaCl in distilled water, freshly prepared.

II.2.6. *C. elegans* RNA interference

IPTG (isopropyl β-D-1-thiogalactopyranoside)

1 M stock solution in distilled water, sterilized by filtration, stored at -20 °C.

RNAi feeding agar plates

50 µg/ml ampicillin, 1 mM IPTG in NGM agar, stored at 4 °C, protected from light exposure.

II.2.7. *C. elegans* embryo immunostaining

PBS (phosphate-buffered saline) buffer

137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 in distilled water, pH adjusted to 7.4, sterilized by autoclaving.

PBST (PBS Tween-20) buffer

0.05% [v/v] Tween-20 in PBS buffer

Blocking solution

0.5% [w/v] bovine serum albumin, 0.5% [w/v] non-fat milk powder in PBST buffer

II.3. Kits and consumables

Cloning, transformation and DNA purification

Q5 high-fidelity DNA polymerase for cloning PCR (New England BioLabs, Ipswich, USA)

Q5 high GC enhancer (New England BioLabs)

Gateway BP Clonase II Enzyme mix (Life Technologies)

Gateway LR Clonase II Enzyme mix (Life Technologies)

Proteinase K solution (2 µg/µl, Life Technologies)

Mach1 T1 phage-resistant chemically competent *E. coli* for transformation (Life Technologies)

SOC medium (super optimal broth with catabolite repression, Life Technologies)

NucleoBond Xtra Midi plasmid DNA purification kit (MACHEREY-NAGEL, Düren, Germany)

Invisorb Spin Plasmid Mini Two plasmid DNA purification kit (STRATEC Molecular, Berlin, Germany)

Invisorb Spin DNA extraction kit (STRATEC Molecular)

II.4. *C. elegans* cultivation

II.4.1. *C. elegans* strains

C. elegans strains were obtained from Caenorhabditis Genetics Center (CGC, University of Minnesota), Shohei Mitani (National Bioresource Project for the nematode, Tokyo, Japan), Miyeko Mana/Fabio Piano (New York University) and Nikolaus Rajewsky (MDC, Berlin).

Table II.1 | *C. elegans* strains and mutants used in the present thesis.

Strain	Reference
Bristol N2 (wild-type)	
gei-12(tm4526)	
gei-12(tm4259)	
C36C9.1(tm4343)	
pptr-1(tm3103)	(Gallo et al., 2010)
BS1080 *ozIs5[GLD-1::GFP::FLAG; unc-119(+)]*	(Jungkamp et al., 2011)
EK244 *cmIs6[MBK-1::GFP; unc-4(+)]*	(Raich et al., 2003)
JH1576 *axIs1140[Ppie-1::GFP::MBK-2; unc-119(+)]*	(Pellettieri et al., 2003)
JH2015 *axIs1462[Ppie-1::GFP::PIE-1::3'pie-1; unc-119(+)]*	(Merritt et al., 2008)
JH2017 *axIs1464[Ppie-1::GFP::PGL-3::3'pgl-3; unc-119(+)]*	(Merritt et al., 2008)
JH2166 *axIs1567[Ppie-1::GFP::SPN-4::3'spn-4; unc-119(+)]*	(Merritt et al., 2008)
JH2688 *axIs1927[Ppie-1::LAP::GLH-1::3'nos-2; unc-119(+)]*	(Voronina and Seydoux, 2010)
MG170 *(zen-4(or153ts) IV; xsEx6[ZEN-4::GFP])*	(Kaitna et al., 2000)
PF633 *nnIs265[Ppie-1::PAR-6::GFP::3'pie-1; unc-119(+)]*	
PF720 *nnIs352[Ppie-1::POS-1::GFP::3'pie-1; unc-119(+)]*	
PF1207 *nnIs795[Pgld-1::GEI-12::GFP::3xFLAG::3'unc-54; unc-119(+)]*	
SS747 *bnIs1[Ppie-1::GFP::PGL-1; unc-119(+)]*	(Cheeks et al., 2004)
TH120 (*[Ppie-1::GFP::PAR-2::3'par-2]; [Ppie-1::mCherry::PAR-6::3'pie-1]*)	(Schonegg et al., 2007)
WH346 *ojIs34[GFP::CAR-1; unc-119(+)]*	

II.4.2. *C. elegans* culture

C. elegans strains were grown on OP50 *E. coli* bacteria seeded NGM plates using standard techniques as previously described (Brenner, 1974). Worms were maintained at 20°C and transferred to fresh plates twice per week.

II.4.3. SILAC worm culture

SILAC worm culture was performed as reported (Stoeckius et al., 2014). Briefly, peptone-free NGM plates containing 1X antibiotic-antimycotic (Life Technologies) were seeded with concentrated culture of a lysine-auxotrophic *E. coli* strain AT713 (CGSC: 4529, Coli Genetic Stock Center, Yale University), which was metabolically labeled with either Lys-0: L-lysine-$^{12}C_6{}^{14}N_2$ ("light") or Lys-8: L-lysine-$^{13}C_6{}^{15}N_2$ ("heavy") in chemically defined minimal media. Synchronized L1 larvae were added to the seeded plates and grown for one generation. Larvae-to-food (number of larvae versus milliliters of SILAC bacterial culture before being concentrated) ratio was 200:1. An excess of food was maintained during the entire period of cultivation.

II.5. Mammalian standard and SILAC cell culture

HEK293T cells (ATCC, Manassas, USA) were maintained in a humidified incubator under 5% CO_2 at 37°C in DMEM medium (4.5 g/l glucose) supplemented with 10% fetal bovine serum (PAA Laboratories, Cölbe, Germany) and 4 mM GlutaMAX (Life Technologies). Cells were passaged twice per week.

SILAC cell culture was performed using SILAC DMEM media (4.5 g/l glucose) containing ("light") Lys-0: L-lysine-$^{12}C_6{}^{14}N_2$, Arg-0: L-arginine-$^{12}C_6{}^{14}N_4$, or ("medium") Lys-4: L-lysine-$^{12}C_6{}^{14}N_2{}^{2}D_4$, Arg-6: L-arginine-$^{13}C_6{}^{14}N_4$, or ("heavy") Lys-8: L-lysine-$^{13}C_6{}^{15}N_2$, Arg-10: L-arginine-$^{13}C_6{}^{15}N_4$ and supplemented with 10% dialyzed fetal bovine serum (PAA Laboratories), 4 mM GlutaMAX and 1 mM sodium pyruvate. Cells were split every two to three days.

II.6. Cloning

II.6.1. PCR

Polymerase chain reaction (PCR) was performed using manual hot-start technique (polymerase added at the initial denaturation step) in a 50-µl reaction volume containing 1X Q5 reaction buffer, 1 ng of DNA template, 200 µM (each nucleotide) dNTP mix (New England BioLabs), 0.5 µM forward primer (synthesized by BioTeZ,

Berlin), 0.5 µM reverse primer, 0.01 U/µl Q5 high-fidelity DNA polymerase, nuclease-free water and with or without 30% Q5 High GC Enhancer.

The PCR reaction was then carried out using the thermocycling profile as follow: initial denaturation at 98°C for 5 min; 35 cycles of denaturation at 98°C for 20 s, annealing for 30 s (typically, annealing temperature at 60°C for the first 7 cycles and 65°C for the remaining 28 cycles), extension at 72°C for 30-40 s per kilo base pairs; final extension step at 72°C for 5 min.

PCR products were resolved by electrophoresis at constant voltage (100 V) in 1% agarose gel containing ethidium bromide in TAE buffer. The gel was exposed to UV transillumination, and the gel band of expected size was excised with a clean scalpel, followed by purification of the DNA using a commercial gel extraction kit (STRATEC Molecular).

II.6.2. Gateway recombination reaction

Gateway BP reaction was performed in a 5-µl reaction volume containing 25 femtomoles of PCR products, 75 ng of pDONR221 vector, 1 µl of Gateway BP Clonase II Enzyme mix and nuclease-free water. The reaction mix was incubated at room temperature overnight. After that, 1 µg of proteinase K was added, followed by 10-min incubation at 37°C.

Gateway LR reaction was carried out in a 5-µl reaction volume containing 25 ng of entry clone, 50 ng of pDEST destination vector, 1 µl of Gateway LR Clonase II Enzyme mix and nuclease-free water. The reaction mix was incubated at room temperature for 2 hr to overnight. Following that, 1 µg of proteinase K was added and incubated at 37°C for 10 min.

II.6.3. Transformation and plasmid DNA purification

Transformation was performed using a standard heat-shock protocol in Mach1 *E. coli* competent cells. Briefly, 2 µl of Gateway reaction products were added to 50 µl of competent cells and incubated at 4°C for 30 min. Following heat-shock at 42°C for 30 s and incubation at 4°C for 5 min, 250 µl of pre-warmed SOC medium was added and incubated on an orbital shaker at 37°C for 1 hr. The culture was then inoculated using a Drigalski spatula onto LB agar plates containing corresponding antibiotics for selection and incubated

at 37°C overnight. Afterwards, single colonies were inoculated into liquid LB media with the same selection antibiotics and cultivated at 37°C overnight in a shaker incubator.

Plasmid DNA purification was carried out using commercially supplied kits. Briefly, the overnight liquid culture was pelleted by centrifugation at 4°C. The concentrated bacteria were then lysed under alkaline conditions, followed by neutralization. The plasmid DNA was purified by anion-exchange resin and eluted in nuclease-free water. The concentration and purity of the eluted plasmid DNA were then determined by spectrophotometric measurements of absorbance at 230, 260 and 280 nm wavelengths.

II.6.4. DNA sequencing

All cloned sequences in plasmid DNA were verified by Sanger sequencing using external services offered by Source BioScience (Berlin, Germany). Sequencing trace files were carefully checked for sequencing quality and mutations using freely available ApE plasmid editor software.

II.6.5. Construct generation

Full-length *gei-12* was amplified by PCR from N2 cDNA using the Gateway forward primer 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTGAGTTCCTCAAAACCTTACCCA-3' and reverse primer 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTATTGATCTCTGGGTGGGTCAAAAATAG-3', cloned into the pDONR221 vector and subcloned into pCFJ150 vector (Zeiser et al., 2011) for microparticle bombardment (performed by Desirea Mecenas, New York University) and into a pDEST_EGFP destination vector (gift from Markus Landthaler, MDC, Berlin) for mammalian cell expression. Fragments of *gei-12* were further cloned from *gei-12* cDNA using Gateway primers (Table II.2).

For generating the *mbk-2* RNAi clone, a 1393-bp genomic fragment of *mbk-2* was PCR amplified using forward primer: 5'-CGATCACACACATCCTCGTC-3' and reverse primer: 5'-AACCTCATGATCGGCAAGTC-3', TA cloned into the RNAi feeding vector (L4440) and transformed into HT115 bacteria as described (Fraser et al., 2000) (performed by Miyeko Mana, New York University).

Table II.2 | Gateway primers for cloning of *gei-12* fragments.

Fragment	Residue begin	Residue end	Forward: leading sequence[1] + (5' -> 3')	Reverse: leading sequence[2] + (5' -> 3')
N	2	452	AGTTCCTCAAAACCTTACCCAAGCGG	CTTGATCGACAGTGAAGTTTTCTTGGG
N12	2	303	AGTTCCTCAAAACCTTACCCAAGCGG	ACGATCCTGGAAGTTGCTGCTA
N23	144	452	TTCCCAGTATCCACAGATCAGGAC	CTTGATCGACAGTGAAGTTTTCTTGGG
N1	2	143	AGTTCCTCAAAACCTTACCCAAGCGG	TCTTCGTTCATACCAGCTGGAGCTA
N2	144	303	TTCCCAGTATCCACAGATCAGGAC	ACGATCCTGGAAGTTGCTGCTA
N3	304	452	ACTTACTTCAACACAAATGACGATG	CTTGATCGACAGTGAAGTTTTCTTGGG
C	453	862	CGCCGTGGAAGTTCTCGTAGTGCCTC	TTGATCTCTGGGTGGGTCAAAAATAG
C12	453	733	CGCCGTGGAAGTTCTCGTAGTGCCTC	TCTTTCAGGTGACATCCCGTGAAT
C23	593	862	GAAACAGCTTCAGGAAAGCGAATTGCTC	TTGATCTCTGGGTGGGTCAAAAATAG

[1]5'– GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTG
[2]5'– GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA TCA

II.7. Cell biology work

II.7.1. RNAi and sterility assays

Unless stated otherwise, all RNAi clones were from the Ahringer RNAi library (Fraser et al., 2000). Feeding RNAi was performed similar to previously described (Kamath et al., 2001). Individual RNAi clones were streaked onto LB agar plates containing 50 µg/ml ampicillin and 50 µg/ml tetracycline and incubated overnight at 37°C. Single colony was then picked, and the inoculated RNAi bacterial culture was grown in LB media containing 50 µg/ml ampicillin for ~7 hr, followed by IPTG (1 mM) induction for 1 hr. The bacterial culture was then seeded onto NGM plates containing 50 µg/ml ampicillin and 1 mM IPTG. Seeded plates were let dry, protected from light and incubated at room temperature overnight. Synchronized L1 larvae were then added to the seeded plates and were incubated at 25°C for ~48 hr until F1 embryos could be harvested for further examinations.

For trans-generational feeding and sterility assays (performed by Patricia Cipriani, New York University), L1 larvae were subjected to *gei-12(RNAi)* continuously through adulthood on solid medium at 15°C or 25°C; multiple L1 progeny were individually transferred to new feeding plates and the process was repeated through two filial generations. Adults from each generation were examined individually and scored as sterile if no embryos were visible in the uterus.

II.7.2. Mammalian cell transfection

For pull-down experiments, cells were seeded to 15-cm cell culture Petri dishes and grown to 50% confluence. Plasmid DNA (15 µg) and 30 µg of linear polyethylenimine (PEI) transfection reagent (Polysciences, Warrington, USA) were diluted separately in 1.2 ml of serum-free DMEM media. The two solutions were then combined and incubated at room temperature for 30 min. Next, 2.4 ml of serum-free DMEM media were added to a final volume of 4.8 ml. Following a wash with PBS, the transfection mix was added to cells containing 10.2 ml of normal cell culture DMEM media and incubated at 37°C for 24 hr.

For fluorescence microscopy, cells were seeded onto sterile round coverslips (diameter: 18 mm) in a 12-well cell culture plate and grown to 80% confluence. Plasmid DNA (1.6 µg) and Lipofectamine 2000 (4 µl, Life Technologies) were diluted separately in 100 µl of serum-free DMEM media. The two solutions were then mixed together and incubated at room temperature for 5 min. Next, 800 µl of normal cell cluture DMEM media were added to a final volume of 1 ml. Following a wash in PBS, the transfection mix was added to cells and incubated at 37°C for 4 hr. Afterwards, transfection was stopped by replacing the transfection mix with 1 ml of normal DMEM media. The next day, cells were fixed directly without an initial washing step.

II.7.3. Embryo immunostaining and mammalian cell fixation

Young gravid hermaphrodites were dissected to release embryos to a drop of water (7 µl) on a square coverslip. The coverslip was then attached to a glass slide coated with poly-L-lysine, transferred into liquid nitrogen and incubated for at least 10 min. Embryonic cells were exposed by a quick flick of the coverslip using a razor blade to “crack” open the eggshell, followed by an immediate transfer to pure methanol for fixation at -20°C for 10 min. Next, the slide was transferred to pure acetone and incubated at -20°C for 5 min. Rehydration was performed in a descending acetone series (90%, 75%, 50% and 30%) at 4°C modified from Takeda et al. (2008). Following a wash in PBST for 5 min, 25 µl of blocking solution was added to the embryos and incubated in a humidified chamber at room temperature for 1 hr. Afterwards, to stain P granule proteins, blocking solution was replaced by 25 µl of monoclonal K76 antibody solution (1:5 dilution in blocking solution) and incubated at 4°C overnight (Strome and Wood, 1983). After three washes in PBST, 25 µl of FITC-conjugated anti-mouse IgG (1:200 in blocking solution, Jackson ImmunoResearch, West Grove,

USA) was added and incubated at room temperature in the dark for 1 hr. Following another three washes in PBST in the dark, slides were mounted using VECTASHIELD mounting medium with DAPI (Vector Laboratories, Burlingame, USA).

Transfected HEK293T cells grown on coverslips were fixed by 4% paraformaldehyde for 15 min and washed twice in PBS. Following that, the fixed cells were stained with DAPI (1 µg/ml in PBS) for 2 min in the dark and again washed twice in PBS. Slides were mounted using VECTASHIELD HardSet mounting medium (Vector Laboratories).

II.7.4. Fluorescence microscopy

Live and fixed embryo fluorescence imaging and time-lapse microscopy were carried out using 40X or 63X objectives on a Leica DM RA2 microscope equipped with a Hamamatsu C9100-12 EM-CCD camera. Fixed mammalian cell fluorescence images were acquired at 63X with a Zeiss Axio Imager M2 system. Images were processed in Volocity (version 6.1.1, PerkinElmer) or ImageJ software (version 1.49g) (Schneider et al., 2012).

II.8. Quantitative proteomics and affinity purification assays

II.8.1. Benchmark for label-free quantification of whole proteomes

To assess the accuracy of label-free quantification, two label-free samples were prepared in 6M urea/2 M thiourea denaturation buffer: sample 1: 1X MCF-7 (human breast cancer cell line, obtained from Wei Chen, MDC, Berlin) whole cell lysates (10 µg), 1X *E. coli* lysate (5 µg); sample 2: 1X MCF-7 whole cell lysates (10 µg), 4X *E. coli* lysate (20 µg). These two samples were digested in-solution and measured in succession by LC-MS.

II.8.2. Simulation of label-free pull-down experiments

To simulate the detection of label-free pull-down enrichment in a complex non-specific background, two samples of predefined composition were prepared: sample 1: 1X *E. coli* lysate (0.8 µg), 1X UPS2 human standard (0.4 µg, Sigma-Aldrich), recombinant CDC42 (0.2 µg, gift from Florian Paul); sample 2: 1X *E. coli*

lysate (0.8 µg), 4X UPS2 human standard (1.6 µg), recombinant RAC1 (0.2 µg), RHOA (0.1 µg) and GFP (0.4 µg). These two samples were digested in-solution and measured in succession in triplicate by LC-MS.

II.8.3. Embryo pull-down assays

Early embryos were freshly harvested in biological triplicate (~2 million embryos each replicate) by bleaching young gravid hermaphrodites and sonicated on ice (cycle: 0.5 s, amplitude: 40-45%, 5 strokes/session, 5 sessions, interval between sessions: 30 s; UP200S ultrasonic processor (Hielscher Ultrasonics, Teltow, Germany)) in lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 10% glycerol, protease inhibitor cocktail (Roche, Basel, Switzerland), 0.1% Nonidet P-40 Substitute (Sigma-Aldrich)). After sonication, Nonidet P-40 Substitute was added up to 1% and the lysates were incubated with head over tail rotation for 30 min at 4°C, followed by centrifugation at 20,000 × g for 20 min at 4°C. Cleared lysate was then aspirated without disturbing the upper lipid layer and split by half into either the anti-GFP agarose beads or the blocked control beads (40-50 µl, ChromoTek, Planegg, Germany). After head over tail rotation at 4°C for 60-90 min, the beads were washed once with lysis buffer containing 0.1% Nonidet P-40 Substitute, followed by two times of washing in either buffer I (25 mM Tris-HCl pH 7.4, 300 mM NaCl, 1 mM $MgCl_2$) or buffer II (1 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$) or both. For GFP::MBK-2 pull-downs, two separate experiments were performed using different washing conditions. Proteins were eluted by orbital shaking in 50 µl of denaturation buffer (6 M urea/2 M thiourea in 10 mM HEPES pH 8.0) at room temperature without an ethanol precipitation step. For the MBK-1::GFP pull-down experiments, proteins were eluted twice by shaking in 50 µl of 8 M guanidinium chloride at 90 °C, followed by ethanol precipitation.

II.8.4. Pull-downs using SILAC worms to assess post-lysis bindings

For checking the impact of post-lysis interactions, SILAC-labeled “light” BS1080 (GLD-1::GFP) young adults were mixed with “heavy” N2 worms, either before lysis or only at the last washing step before elution (Fig. III.9 A). For the samples mixed before the pull-down, lysates were incubated with anti-GFP agarose beads for 60 min and the bound proteins were eluted in denaturation buffer. For the samples mixed after the pull-

down, the incubation time was 30 min. Elution was performed three times (100 µl each) in 100 mM glycine-HCl pH 2.5 and neutralized by adding 10 µl of 1 M Tris-HCl pH 8.0, followed by ethanol precipitation.

II.8.5. GEI-12 pull-downs in mammalian cells

For the SILAC pull-down experiments using HEK293T cells, the EGFP::GEI-12 (full-length or fragments) expressing cells and the control EGFP-only expressing cells were lysed separately with a Dounce homogenizer in lysis buffer (25 mM Tris-HCl pH 7.4, 125 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 5% glycerol, protease inhibitor cocktail, 1% Triton X-100). Lysates were cleared by centrifugation (20,000 × g for 20 min at 4°C) and then incubated with anti-GFP agarose beads at 4°C for 90 min, followed by three sequential washes in the following buffers: I (25 mM Tris-HCl, pH 7.4, 125 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA, 0.1% Triton X-100), II (25 mM Tris-HCl pH 7.4, 125 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA), III (1 mM Tris-HCl pH 7.4, 150 mM KCl, 1 mM $MgCl_2$). Beads of the individual SILAC states were combined before the final wash. Proteins were eluted using 8 M guanidinium chloride at 90 °C and precipitated by ethanol.

To check whether GEI-12 interactions were partially mediated by RNA, another pull-down experiment against EGFP::GEI-12 was performed using cell lysates pre-treated with or without nuclease (250 U, Pierce Universal Nuclease, Thermo Scientific) for 20 min.

II.9. Ethanol protein precipitation and in-solution digestion

Protein samples that had been eluted by glycine-HCl or guanidinium chloride were precipitated by ethanol. For glycine-HCl eluted samples, an additional 70 µl of 2.5 M sodium acetate (pH 5.0) was added. Sample volume was filled up to 2 ml with pure ethanol, followed by addition of 2 µl of Glycoblue (Life Technologies). Next, the sample was incubated at 4°C overnight. Afterwards, proteins were pelleted by centrifugation at 20,000 × g for 60 min at 4°C. Supernatant was carefully decanted, and the protein pellet was air-dried and resolubilized in 30 µl of denaturation buffer (6 M urea/2 M thiourea in 10 mM HEPES pH 8.0).

The protein solution in denaturation buffer was reduced in 10 mM DTT for 40 min and alkylated in 5.5 mM iodoacetamide in the dark at room temperature for 30 min. After that, protein digestion was started by

adding endoproteinase Lys-C (lysyl endopeptidase, Wako, Osaka, Japan) at a protein:enzyme ratio of 50:1, followed by 3-hr incubation at room temperature with orbital shaking. Next, the sample was diluted 1:4 with 50 mM ABC buffer and further digested by adding trypsin (sequence grade, Promega, Madison, USA) at a protein:enzyme ratio of 50:1. Of note, the trypsin digestion step was skipped for SILAC worm pull-down samples that had been labeled with "heavy" Lys-8 only. The sample was then incubated at room temperature with orbital shaking overnight. Afterwards, digestion was terminated by 0.7% [v/v] trifluoroacetic acid.

II.10. StageTip purification

Digested peptides were desalted and concentrated by StageTip (stop and go extraction tip) containing 3 disks of C_{18} reverse-phase material (Empore, 3M, Minneapolis, USA) (Rappsilber et al., 2007). Briefly, the StageTip was activated by 80 µl of pure methanol followed by a wash using 100 µl of buffer A*. Next, the peptide solution was loaded to the StageTip and centrifuged at 2,500 × g, followed by another wash using 100 µl of buffer A*. The StageTip was then stored at 4°C until elution. For LC-MS measurement, the purified peptides were eluted by 60 µl of buffer B to a microtiter plate and dried by vacuum centrifugation to a volume of 2~3 µl. Afterwards, 5 µl of buffer A was added to each sample.

II.11. Liquid chromatography tandem mass spectrometry

Peptide mixtures were separated by reversed phase chromatography using the Eksigent NanoLC Ultra system or the EASY-nLC system (Thermo Scientific) on a 20-cm fritless silica microcolumn (inner diameter: 75 µm) packed in-house using ReproSil-Pur C18-AQ 3-µm resin (Dr. Maisch GmbH, Ammerbuch, Germany). Peptides were separated on an 8-50% acetonitrile gradient (Velos: 200 min; Q Exactive: 120 min) with 0.5% formic acid at a nanoflow rate of 200 nl/min. Eluting peptides were directly ionized by electrospray ionization at 2.2 kV and transferred into an LTQ Orbitrap Velos or a Q Exactive Quadrupole-Orbitrap hybrid mass spectrometer (Thermo Scientific).

For all pull-down experiments, mass spectrometry was conducted using data-dependent mode with one full scan (MS) followed by fragmentation scans (MS/MS) of the 20 most intense ions (Velos) or 10 most intense ions (Q Exactive). MS scan was performed in the Orbitrap mass analyzer with the following settings: for Velos, m/z (mass-to-charge ratio) range: 300-1,700, resolution: 60,000, target value: 10^6; for Q Exactive, m/z range: 300-1,700 or 375-1,700, resolution: 70,000, target value: 3×10^6, maximum injection time: 120 ms. MS/MS scan was conducted using the following settings: for Velos, collision-induced dissociation (CID) in the linear trap quadrupole (LTQ), target value: 3×10^3, isolation width: (+/-) 2.0 m/z, normalized collision energy: 40%, dynamic exclusion time: 60 s, monoisotopic precursor selection enabled, wideband activation enabled; for Q Exactive, higher energy collision dissociation (HCD), resolution: 35,000, target value: 5×10^5, maximum injection time: 120 ms, isolation window: 4.0 m/z, normalized collision energy: 26%, dynamic exclusion time: 30 s. Ions with an unassigned or +1 charge state were rejected for precursor ion isolation.

For the simulation experiments, mass spectrometry was performed on the Q Exactive instrument with the following settings: for peptide separation: 200-min acetonitrile gradient (8-50%); for the MS scan: m/z range: 300-1,700, resolution: 70,000, target value: 1×10^6, maximum injection time: 120 ms; for the MS/MS scans using HCD: resolution: 17,500, target value: 5×10^5, maximum injection time: 60 ms, isolation window: 2.0 m/z, normalized collision energy: 26%, dynamic exclusion time: 30 s.

II.12. Mass spectrometry data analysis

Raw data files were processed by MaxQuant software (version 1.4.1.2) (Cox and Mann, 2008) using the built-in Andromeda search engine against a target-decoy database containing the forward and reverse sequences: for *C. elegans* pull-down data, WormPep release WS245 and *E. coli* K-12 MG1655 proteome; for human cells, Uniprot human proteome release 201402. By default, sequences of 247 common contaminants, such as keratins, were also included in the search. Carbamidomethylation of cysteine was used as fixed modification; oxidation of methionine and acetylation of the protein N-terminus were set as variable modifications. Trypsin/P specificity (cleave C-terminal to lysine and arginine residues also if a proline follows) was chosen

for all data except the SILAC worm data, for which LysC/P specificity was used (cleave C-terminal to lysine also if a proline follows). Minimal peptide length of 7 amino acids was required and a maximum of 2 missed cleavages were allowed. The "second peptide" option was chosen to decipher co-fragmented MS/MS spectra. False discovery rate (FDR) was set to 1% for both peptide and protein identifications. Each protein group was required to contain at least one unique peptide.

For *in vivo* label-free pull-down experiments, protein quantification was performed using the label-free quantification (LFQ) algorithm (Cox et al., 2014). Minimum LFQ ratio count was set to one. Both unique and razor peptides were considered for quantification. Retention times were recalibrated using the default nonlinear time-rescaling algorithm. The "match between runs" option was chosen for transferring MS/MS identifications between LC-MS runs with the maximal retention time window set to 1 min. Only proteins quantified in at least two out of the three GFP pull-down replicates were included in the analysis. LFQ intensities were $\log_2$-transformed, and imputation for missing values was performed in Perseus software (version 1.2.0.17) (Cox and Mann, 2012) based on a simulated normal distribution to represent low abundance values below the noise level (width = 0.3; shift = 1.8). The LFQ abundance ratio was then calculated for each protein between the GFP pull-downs and the controls. Significance of the enrichment was measured by two-sample Student's *t*-test assuming equal variances. Specific interaction partners were then determined in a volcano plot where a combined threshold (hyperbolic curve) was set based on a modified *t*-statistic (s_0 = 1.5, t_0 = 0.9 ~ 1.1) (Li, 2012; Tusher et al., 2001). Threshold values were chosen to balance sensitivity and false discovery rate based on the simulation experiment with spike-in standards (Fig. III.2). Proteins cross-reactive to the anti-GFP beads were identified by a pull-down assay using wild-type N2 embryos (Fig. III.6). These proteins as well as HDA-3, a protein identified in almost all other pull-downs, have been filtered out.

For SILAC experiments, maximum of 3 labeled amino acids per peptide were allowed. "Requantify" option was chosen. For label-swap SILAC pull-downs, only proteins quantified in both forward and reverse experiments were considered for analysis with a minimum SILAC ratio count set to one. For other SILAC pull-down experiments, a minimum SILAC ratio count of two was required.

II.13. SDS-PAGE and Western blotting

Protein samples were reduced and denatured at 90°C for 5 min in SDS sample loading buffer containing 100 mM DTT. Afterwards, proteins were resolved at constant voltage (100 V) in a NuPAGE precast 4-12% gradient gel (Life Technologies) and transferred to a PVDF blotting membrane (Millipore, Billerica, USA) for 2 hr at constant current (250 mA) in a tank transfer system. Blocking was done in 5% non-fat milk in TBST buffer for 1 hr, followed by primary antibody incubation at 4°C overnight. Primary antibodies were diluted in blocking solution as: 1:500 anti-CAR-1 (Boag et al., 2005), 1:500 anti-IFET-1 (Sengupta et al., 2013) and 1:250 anti-SPN-4 (Huang et al., 2002). After three washes (5 min each) in TBST, the membrane was incubated with species-specific horseradish peroxidase-conjugated secondary antibodies (1:10,000, Sigma-Aldrich) at room temperature for 1 hr, followed by another three washes in TBST. The membrane was then briefly incubated with substrates for enhanced chemiluminescence (Pierce ECL Plus). Signal was detected by a GE Typhoon FLA 9000 biomolecular imager at 473 nm excitation wave-length with an LPB filter.

II.14. Statistical and network analyses

Statistical and network analyses were performed using R (version 3.1.2), Bioconductor and Cytoscape (version 3.1.1). Gene ontology (GO) enrichment analysis for the CAR-1 interactome was conducted using the GOstats package (Falcon and Gentleman, 2007). Other enrichment analyses were done based on the hypergeometric distribution. RNAi and GO annotations were retrieved from WormBase release WS245. Known RNP granule proteins were retrieved from Kato et al. (2012) and converted from mouse to human orthologs using the HomoloGene database (Coordinators, 2015).

For comparison analysis with IVI (*in vivo* interactome), homomeric interactions and interaction pairs containing “dead” genes, which are no longer valid gene model entries in WS245, were removed from the WI8 (Worm Interactome version 8) and LCI (literature-curated interaction) datasets (Simonis et al., 2009). Different alternative splicing isoforms of the same gene were collapsed into one. For enrichment analyses, control gene sets were generated by selecting 3,000 pairs of protein-coding genes at random from WormPep

(WS245). The mean percentage of gene pairs with shared RNAi phenotypes or GO terms across 100 iterations of independent control sets was then used for comparison with the IVI, WI8 and LCI interactome datasets.

Protein disorderness prediction was performed using the IUPred algorithm (Dosztanyi et al., 2005). Prediction of binding regions within the disordered parts of a protein was carried out using the ANCHOR algorithm (Dosztanyi et al., 2009). The EMBOSS software suite (Rice et al., 2000) was used to calculate protein charge distribution, summarized as mean charges of a sliding window of 10 amino acids along the length of the protein.

III Results

III.1. *In vivo* interaction proteomics method development

III.1.1. Technical assessment and optimization of label-free quantification

III.1.1.1 Benchmark of whole proteome label-free quantification

Mass spectrometry is not inherently quantitative (Aebersold and Mann, 2003). Detected peptides are used to infer the identity and quantity of the proteins from which they are derived. However, the ion signals of peptides detected by a mass spectrometer are not direct measurements of the absolute abundances of the proteins present in the original sample. This is due to various reasons, such as unequal ionization efficiencies of different peptides, under-sampling of the instrument, loss of proteins and peptides during sample processing procedures. For relative quantification of multiple samples, stable isotope-based methods using either metabolic labeling or *in vitro* chemical-labeling have become standard practice in the last decade. Stable isotope labeling has the advantage that it allows different samples to be combined for the same LC-MS measurement. Since the labeled and unlabeled peptides are identical in chemical composition, they have nearly equal chromatographic behaviors while remaining distinguishable by the mass spectrometer due to the mass difference.

Although generally considered not as accurate as labeling-based approaches, label-free quantification remains a widely used option. This is mainly due to its simplicity in sample preparation procedures and the unrestricted number of samples that can be compared, particularly when primary tissues are studied. Over the years, a number of different quantification algorithms have been proposed, from simple spectral counting, such as protein abundance index (PAI) and exponentially modified PAI (emPAI) (Ishihama et al., 2005; Rappsilber et al., 2002), to extracted ion chromatogram (XIC)-based approaches, which integrate peptide ion intensities over its chromatographic profile. As expected, XIC-based methods were demonstrated to out-perform spectral counting approaches with regard to the accuracy of quantification (Ahrne et al., 2013; Grossmann et al., 2010).

Recently, the label-free quantification (LFQ) algorithm implemented in the MaxQuant software, has been shown to be an accurate and robust approach for quantifying relative protein abundances (Cox et al., 2014). To benchmark the quantification accuracy of the LFQ algorithm, two label-free samples were prepared: both samples contained equal amounts of MCF-7 (human breast cancer cell line) whole cell lysates; sample 2 contained four times the amount of *E. coli* lysates relative to sample 1. These two samples were measured in succession using both the Velos and Q Exactive Orbitrap instruments. As depicted in Fig. III.1, the two protein populations (human vs. *E. coli*) were well separated by approximately a $\log_2$ fold change of 2, corresponding to the 4-fold difference in the relative abundance of *E. coli* proteins. However, quantification accuracy declined as the LFQ intensity became lower, displaying a Christmas tree-like distribution which is typically observed in mass spectrometry-based protein quantification (Cox and Mann, 2008), suggesting that performing multiple replicate experiments may help further improve the accuracy and robustness of LFQ quantification.

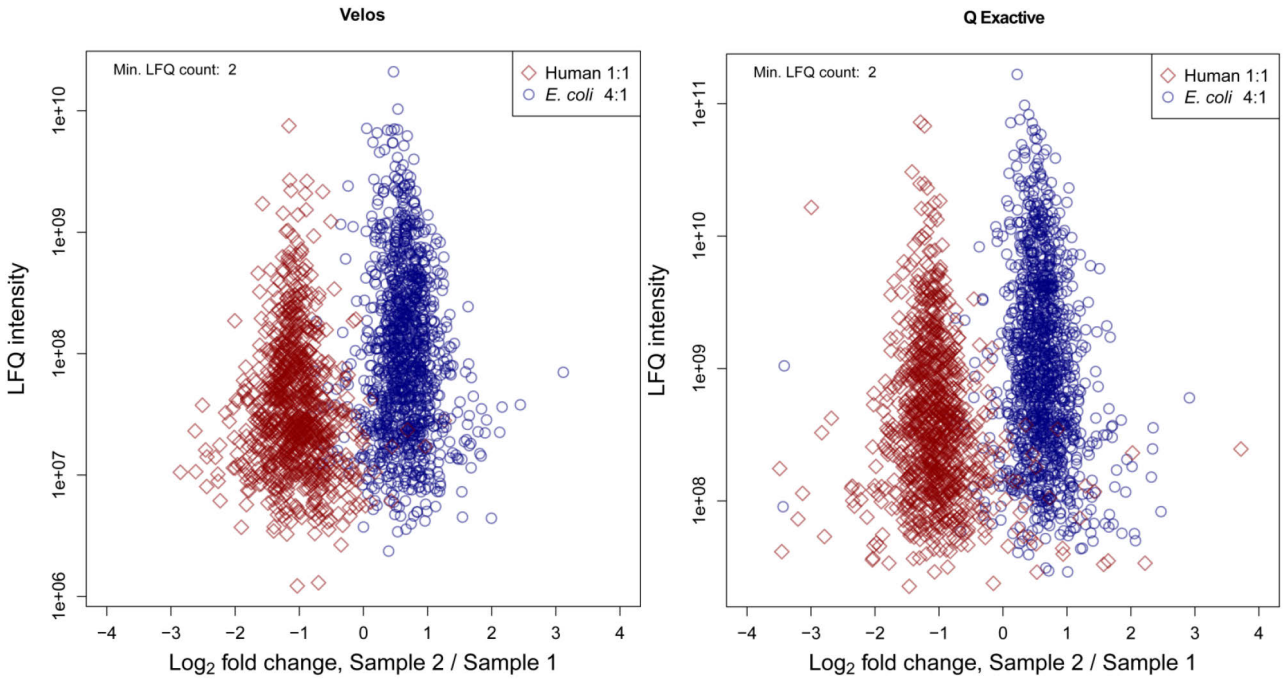

Fig. III.1 | Benchmark of whole proteome label-free quantification. Two label-free samples measured in succession by a Velos or a Q Exactive Orbitrap mass spectrometer. LFQ, label-free quantification.
Sample 1: 1X human MCF-7 cell lysate, 1X *E. coli* lysate;
Sample 2: 1X human MCF-7 cell lysate, 4X *E. coli* lysate.

III.1.1.2 Simulation of label-free pull-down assays

In affinity purification of protein complexes coupled with quantitative mass spectrometry, the distinction between specific interactions and non-specific background binders relies on quantitative information of enriched proteins compared to the negative control (Vermeulen et al., 2008). Modern mass spectrometers are fast and highly sensitive. Nowadays, it is common to detect at least hundreds of non-specific contaminants binding to the solid matrix in a single pull-down assay. Lately, a report of nearly 200 pull-down experiments identified half of the yeast proteome as background binders (Keilhauer et al., 2015). Therefore, to detect a small number of enriched proteins in a complex background mixture remains a challenging task. In recent years, pull-down assays employing multiple biological replicates, label-free quantification and robust statistics have been demonstrated to be an effective approach for detecting specific protein-protein interactions with high confidence, both in cell lines and animal tissues (Andlauer et al., 2014; Hanack et al., 2015; Hubner et al., 2010; Hubner and Mann, 2011).

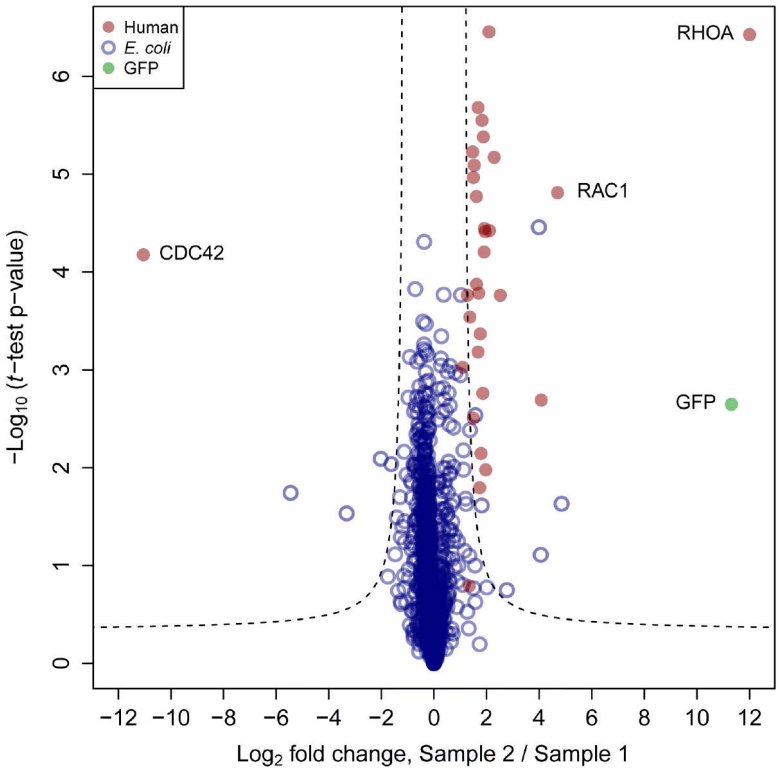

Fig. III.2 | Simulation of label-free pull-down experiments. To simulate the detection of label-free pull-down enrichment in a complex non-specific background, two predefined samples were measured in succession in triplicate. Sample 1: 1X *E. coli* lysate, 1X UPS2 human standard, recombinant CDC42; Sample 2: 1X *E. coli* lysate, 4X UPS2 human standard, recombinant RAC1, RHOA and GFP. A combined threshold (hyperbolic curves) was set based on a modified *t*-statistic (s_0 = 1.5, t_0 = 0.8) (Li, 2012; Tusher et al., 2001).

To simulate pull-down assays in identifying specifically enriched proteins among a complex background mixture and to assess the accuracy of label-free quantification in such an experimental set-up, two samples with predefined protein composition were prepared. These two label-free samples contained the same amount of *E. coli* whole cell lysates (to simulate background binders). In addition, both samples contained a mixture of 48 different isolated or recombinant human proteins (UPS2 standard) in a 1:4 ratio. Finally, four additional recombinant proteins were exclusively spiked-into one of the two samples (Sample 1: CDC42; Sample 2: RAC1, RHOA and GFP). These two samples were measured in succession in triplicate by LC-MS. As displayed in Fig. III.2, 29 out of 32 detected differential binders were correctly identified. In parallel, 1174 out of 1184 background binders (*E. coli*) were also correctly classified. Thus, this benchmark simulation assay had a total accuracy of 98.9% and precision of 90.6%, indicating very high accuracy of the method in distinguishing enriched binders from background contaminants.

III.1.2. Optimization of specificity threshold

III.1.2.1 Definition of the combined cut-off value (hyperbolic curve)

In spite of the quantitative information, drawing a cut-off line to differentiate between specific and non-specific binders is not trivial. With multiple experimental replicates and statistics checking consistency of the enrichment, both the size (the log fold change) and the statistical significance of the enrichment can be assessed in a single volcano plot. This has been well-established practice for large-scale datasets involving multiple replicates, such as microarray data. Typically, two empirical thresholds of fixed values are drawn, as indicated by two crossed lines perpendicular to each other. However, instead of two arbitrary “hard” cut-offs, a widely adopted alternative is to use the SAM (statistical analysis of microarray) algorithm developed by Tusher et al. (2001).

The standard Student’s *t*-statistic is defined as

[Eq1]

$$t = \frac{\log_2 FC}{s}, where\ s\ is\ the\ standard\ error\ of\ the\ difference\ of\ the\ two\ means.$$

The core of the SAM algorithm is essentially a modified Student's *t*-statistic (t_{SAM}). The only difference compared to a standard *t*-statistic (*t*) is the small penalty value (s_0) added to the denominator to stabilize the coefficient of variation across the range of intensities. Effectively, the t_{SAM} value is a signal-to-noise ratio with a corrected noise level (Li, 2012; Tusher et al., 2001).

[Eq2]

$$t_{SAM} = \frac{\log_2 FC}{s + s_0}$$

This allows for defining only a single threshold value of minimum t_{SAM} as t_0 to take both the levels of signal (enrichment) and noise (variation) into account (Li, 2012).

[Eq3]

$$t_{SAM} = \frac{\log_2 FC}{s + s_0} \geq t_0$$

When equal variances are assumed, the degree of freedom for a two-sample *t*-test is fixed, and therefore, there will be a one-to-one correspondence between *t* and the *p* value (Li, 2012). Accordingly, *s* can be substituted by

[Eq4]

$$s = \frac{\log_2 FC}{t}$$

and the t_{SAM} equation [Eq3] can then be transformed into

[Eq5]

$$-\log_{10}(p\ value) \propto t = \frac{|\log_2 FC| \times t_0}{|\log_2 FC| - s_0 \times t_0}, where\ |\log_2 FC| > s_0 \times t_0;\ equal\ variences\ assumed.$$

This is the equation for the combined cut-off (hyperbolic curve) that is defined by a single threshold value of t_0, where the curve approaches $s_0 \times t_0$ on the *x*-axis of the volcano plot and the corresponding *p* value of t_0 on the *y*-axis, as depicted in Fig. III.3.

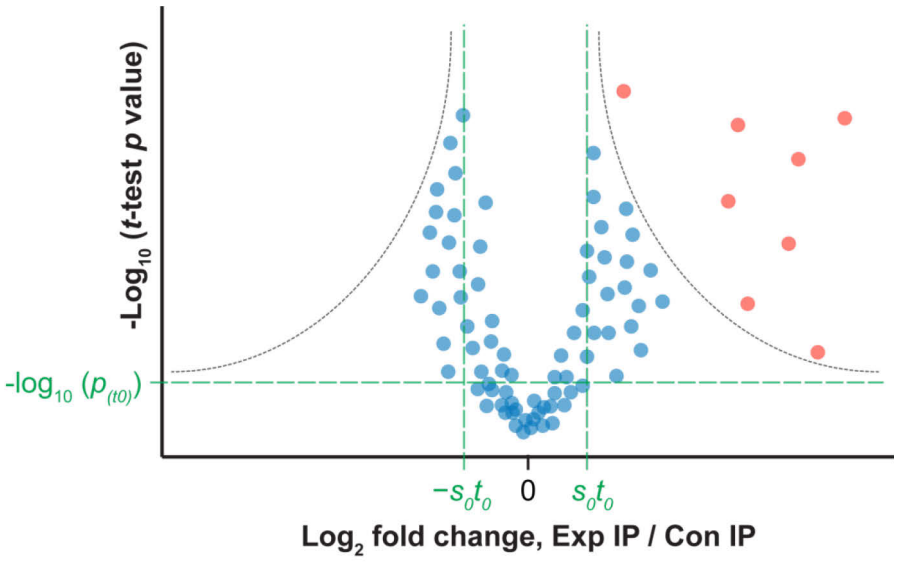

Fig. III.3 | Diagram of a hypothetical volcano plot for label-free quantitative pull-down assays. The dotted line in grey (hyperbolic curve) defines a combined cut-off to distinguish the specific binders (red) from the non-specific background contaminants (blue). The green dashed line depicts the limit that the hyperbolic curve approaches: $s_0 \times t_0$ on the *x*-axis; the corresponding *p* value of t_0 on the *y*-axis. IP, immunoprecipitation.

III.1.2.2 Optimization of the combined cut-off

The current thesis employed an empty-bead negative control approach. Thus, theoretically no proteins should be specifically enriched in the control pull-down. If any proteins appeared to be enriched on the left side of the volcano plot (Fig. III.3), they would be considered false positive hits. In the pull-down simulation data (Fig. III.2), since all the *E coli.* proteins are expected to be categorized as background binders, any detected *E. coli* protein deviated from such is considered a false positive. Therefore, it is possible to calculate the true false discovery rate given a certain threshold value (t_0). Iteration of this procedure over a range of s_0 and t_0 values (varying the hyperbolic curves) showed that the number of false positive hits circumscribed by the hyperbolic curve on the left side of the volcano plot is a good estimation of the false positives on the right side (Fig. III.4). In other words, in the embryo pull-down samples where the distinction between specific and background binders will solely rely on the combined cut-off, minimizing positive hits on the left side of the volcano plot would suggest a low false discovery rate in the identified specific interactions. Further optimization was assisted by iterative calculation of the sensitivity and false discovery rate of the assay with regard to changing thresholds of t_0 (Fig. III.5). Taken all together, the determination of the combined cut-off

values was based on a balance between sensitivity and false discovery rate, and the following values were used for the embryo *in vivo* protein-protein interaction screen data: s_0 = 1.5, t_0 = 0.9 ~ 1.1.

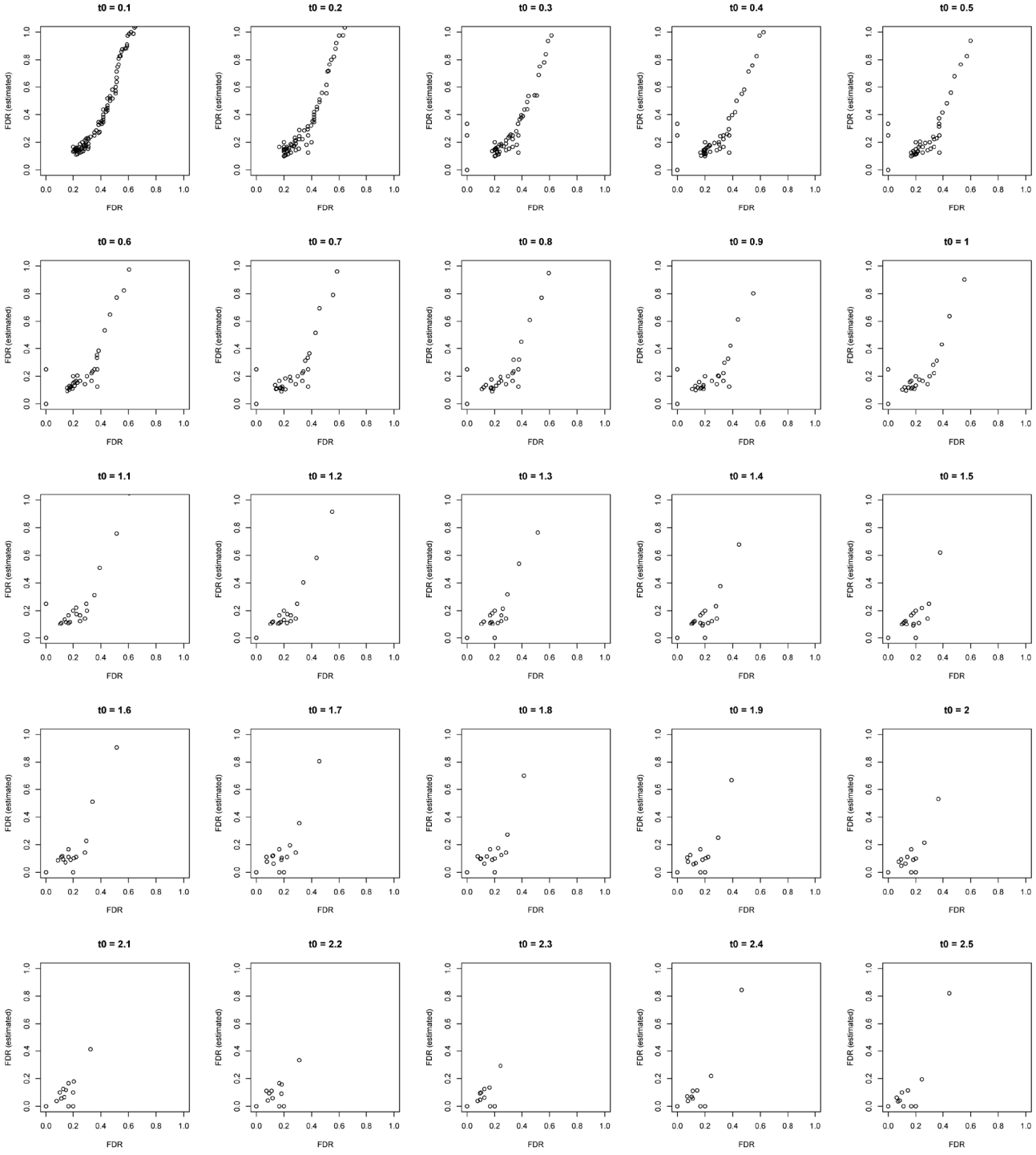

Fig. III.4 | Optimization of parameters for determining the combined cut-off based on pull-down simulation data. FDR on the *x*-axis represents false positives (*E. coli*) circumscribed by the hyperbolic curve on the right side of the volcano plot. FDR (estimation) on the y-axis represents the number of false positives on the left side of the volcano plot divided by the total positive hits on the right side. Iteration over s_0 = 0 ~ 20, t_0 = 0.1 ~ 2.5, interval = 0.1 for both. FDR, false discovery rate.

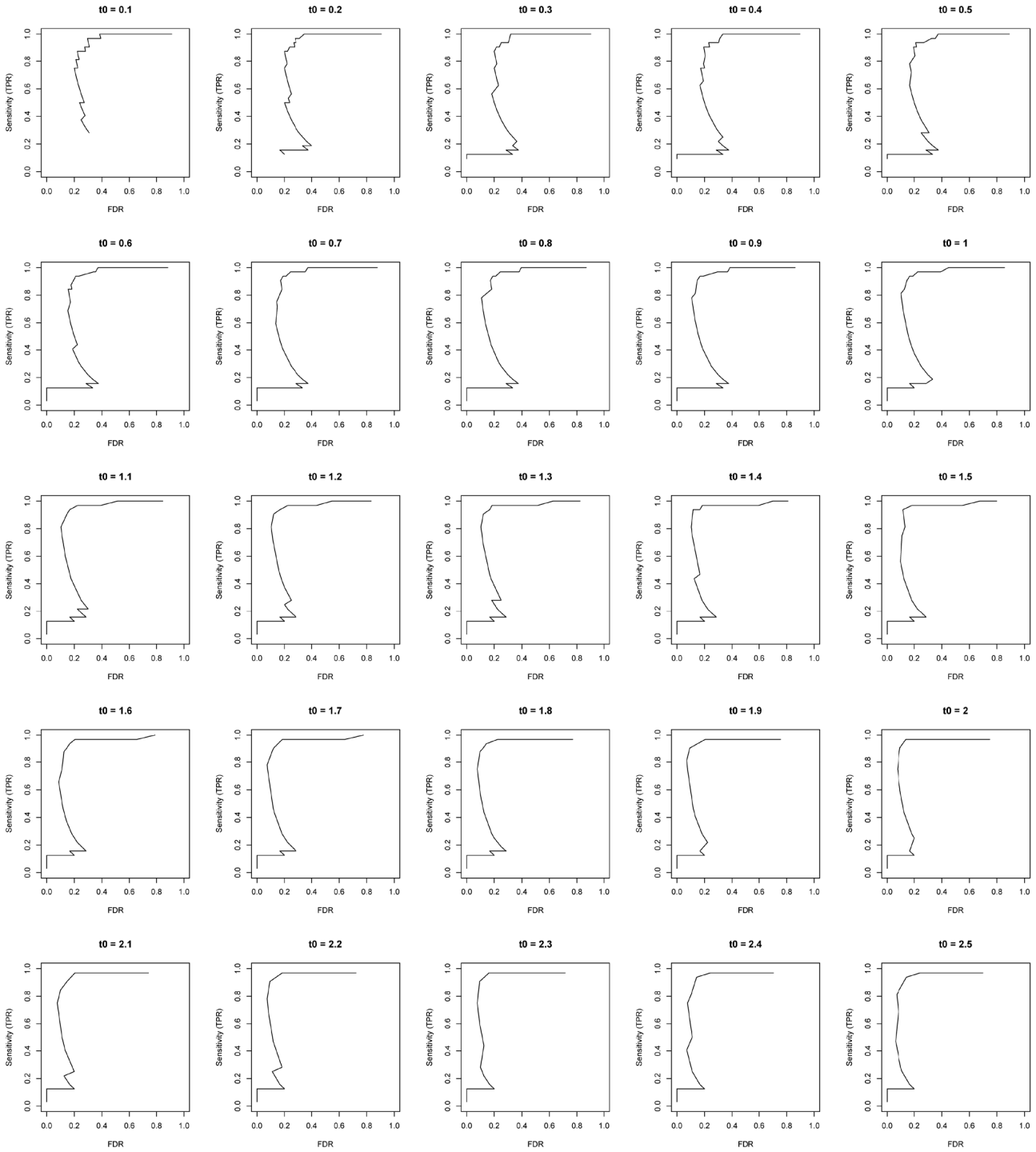

Fig. III.5 | Optimization of parameters for determining the combined cut-off based on pull-down simulation data. FDR on the *x*-axis represents false positives (*E. coli*) circumscribed by the hyperbolic curve on the right side of the volcano plot. Sensitivity (TPR) on the y-axis represents the number of identified true positives divided by the total number of detected theoretical differential binders. Iteration over s_0 = 0 ~ 20, t_0 = 0.1 ~ 2.5, interval = 0.1 for both. FDR, false discovery rate; TPR, true positive rate.

III.1.3. Minimal cross-reactivity of the anti-GFP antibody

As already mentioned in the introduction section, the choice of a proper negative control strategy is critical for quantitative interaction proteomics. The bead-only control approach removes the use of a control strain and thereby completely eliminates any potential biases caused by strain difference or variation in sample preparation steps prior to affinity purification. While this approach offers certain advantages, the prerequisite is that there should be minimal cross-reactivity of the antibody to any protein other than the desired antigen. Pull-down assays using wild-type N2 embryos which did not express GFP showed that the single-chain anti-GFP antibody (Rothbauer et al., 2008) used in the present thesis had a very low cross-reactivity to *C. elegans* proteins (Fig. III.6). These cross-reactive proteins have also been filtered out from the final results of the embryo *in vivo* protein-protein interaction screen.

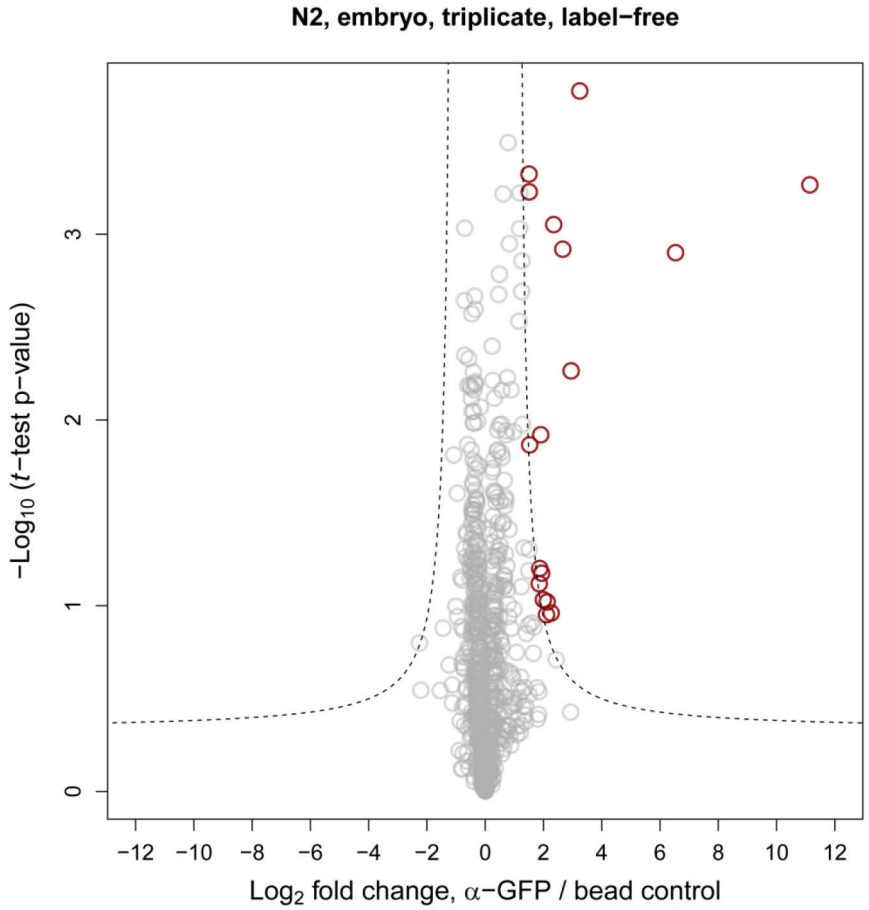

Fig. III.6 | Volcano plot showing that the background binding profiles of control beads and anti-GFP beads are very similar in wild-type N2 worms which do not express GFP fusion proteins. Proteins cross-reactive to the anti-GFP beads (highlighted in red) were filtered out from the results of pull-down assays against GFP fusion proteins.

III.1.4. CAR-1 interactome in early embryos as proof of concept

To assess the feasibility of the label-free quantitative pull-down approach to indentify interaction partners in *C. elegans* early embryos, the bait protein CAR-1 was chosen, whose interaction with the RNA helicase CGH-1 is evolutionarily conserved (Audhya et al., 2005; Boag et al., 2008; Boag et al., 2005), and their interaction partners had been well studied (Boag et al., 2008; Boag et al., 2005; Sengupta et al., 2013). CAR-1 (Cytokinesis, Apoptosis, RNA-associated) is a maternally loaded RNA-binding protein with multiple essential functions during germ cell development and embryogenesis. As its gene name indicates, it has been shown to be required for early embryonic cytokinesis, inhibit physiologic germ cell apoptosis during oocyte development and bind CGH-1 in an RNA-dependent manner (Audhya et al., 2005; Boag et al., 2005). In embryos, CAR-1 localizes to cytoplasmic processing bodies (P bodies) as well as P granules (Audhya et al., 2005; Updike and Strome, 2010).

In the pull-down assays against GFP::CAR-1, about 2 million transgenic early embryos were harvested from young gravid hermaphrodites by bleaching. Three biological replicates were prepared in parallel. For each replicate, the same embryo lysates were split by half into anti-GFP pull-down and the control pull-down. The abundances of the co-purified proteins were assessed by label-free quantitative mass spectrometry to identify proteins specifically enriched in the GFP::CAR-1 pull-downs (Fig. III.7 A). A total number of 429 proteins were identified in the pull-down assays. Combined with robust statistics as shown in a volcano plot (Fig. III.7 B), most of the detected proteins had similar abundances in both the GFP and the control pull-downs, indicating they were non-specific background binders. Importantly, the 34 proteins significantly enriched in the GFP::CAR-1 pull-downs include the bait proteins, GFP and CAR-1, and almost all the proteins (CEY-2, CEY-3, CEY-4, IFET-1, PAB-1) that have been previously found to interact with endogenous CAR-1/CGH-1 *in vivo*, with the exception of ATX-2 and PATR-1 (Boag et al., 2008; Boag et al., 2005; Sengupta et al., 2013). These results show that the assay targeting GFP fusion protein can efficiently identify *in vivo* interaction partners of the endogenous protein with good sensitivity.

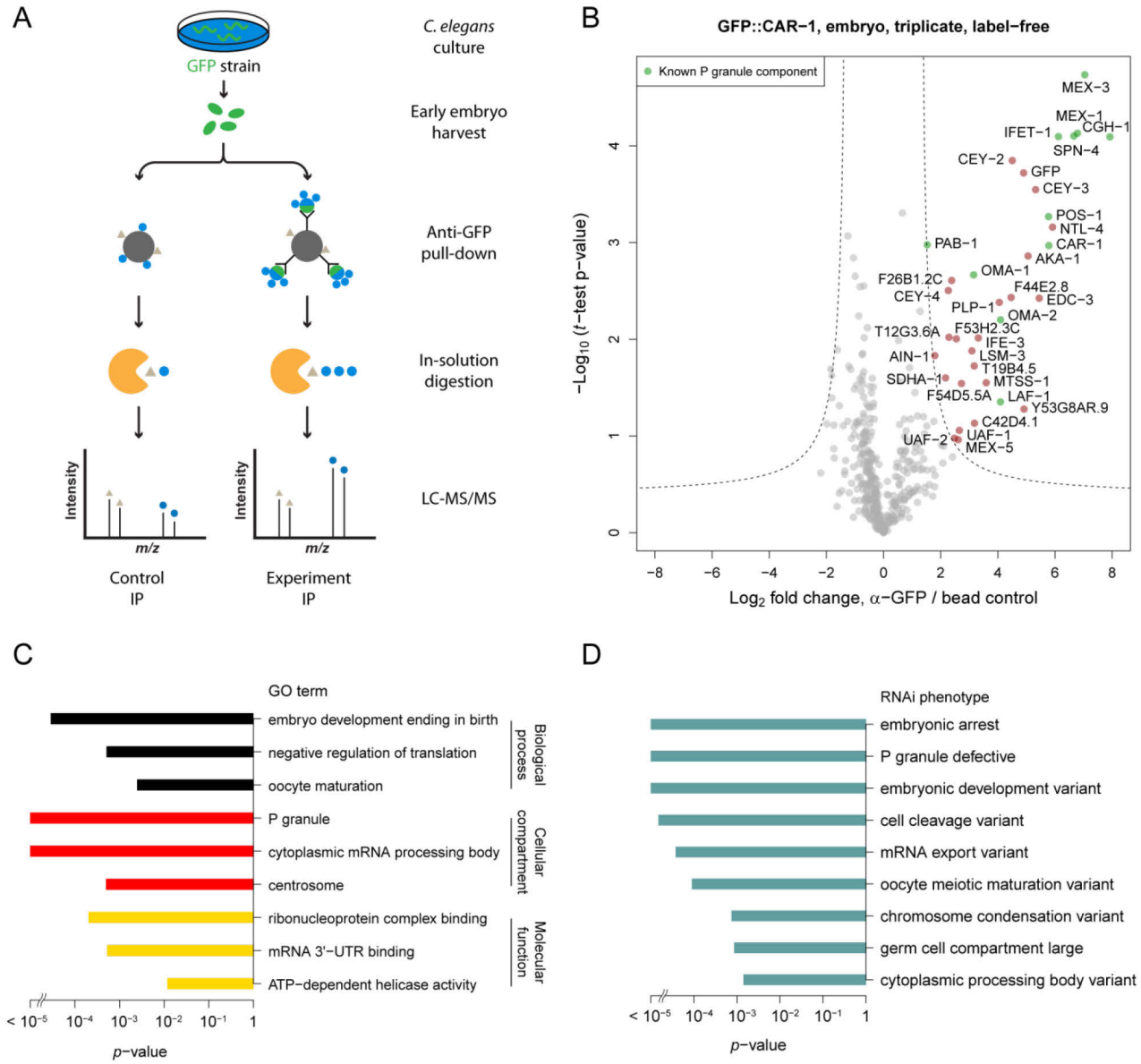

Fig. III.7 | Label-free *in vivo* interaction proteomic strategy and proof-of-concept experiment. (A) Experimental set-up. The same extract of early embryos expressing GFP fusion proteins was split and incubated separately with anti-GFP or control beads. The abundances of co-purified proteins were measured by mass spectrometry. Three biological replicates were performed in parallel. (B) Volcano plot showing specific interaction partners of GFP::CAR-1 based on pull-down experiments of three biological triplicates. Interaction partners meeting stringency thresholds (hyperbolic curves) are shown in color. Known P granule components among the interaction partners are shown in green. (C) GO term enrichment analysis for the interaction partners of GFP::CAR-1. Selected significantly enriched GO terms are displayed (conditional hypergeometric test). (D) RNAi phenotype enrichment analysis for the interaction partners of GFP::CAR-1. Selected significantly enriched RNAi phenotypes are shown (hypergeometric test).

Among the identified interaction partners, many are known to be at least transiently associated with P granules during development (Sengupta et al., 2013; Updike and Strome, 2010). Further gene ontology (GO) analyses revealed that “oocyte maturation”, “P granule” and “cytomplasmic mRNA processing body” were

among the enriched terms annotated for the interaction partners (Fig. III.7 C), consistent with the localization of CAR-1 and its known function during oogenesis (Boag et al., 2005; Squirrell et al., 2006). In parallel, a similar enrichment analysis looking for over-represented RNAi phenotypes revealed that phenotypes such as “embryonic arrest”, “P granule defective”, “cell cleavage variant” and “oocyte meiotic maturation variant” were significantly enriched in the detected CAR-1 interaction partners (Fig. III.7 D). By Western blotting against endogenous proteins following the GFP::CAR-1 pull-down, SPN-4 was confirmed as a novel CAR-1 interaction partner (Fig. III.8). Altogether, these results suggest that the identified *in vivo* interactome can reveal important biological functions of CAR-1.

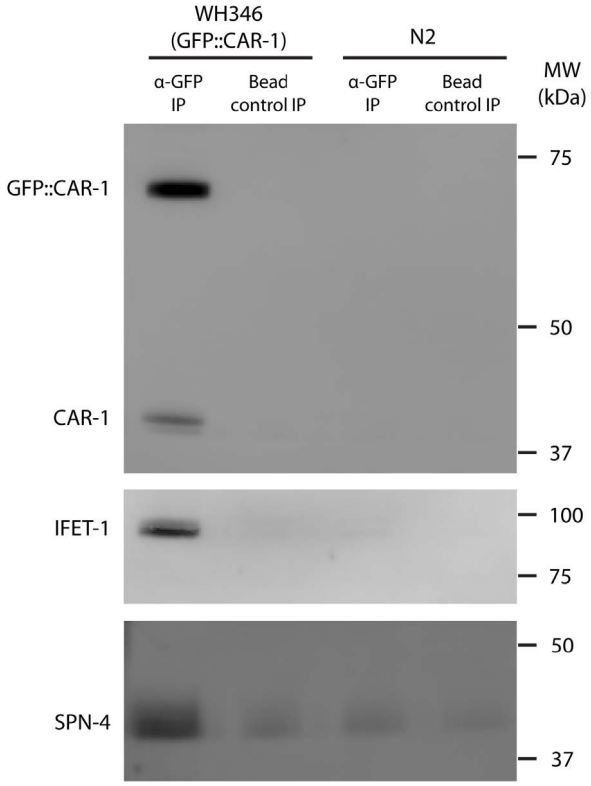

Fig. III.8 | Validation of CAR-1 interaction partners by co-immunoprecipitation and Western blotting. Three examples demonstrating specific binding of GFP::CAR-1 to CAR-1, IFET-1, and SPN-4 proteins in the experimental pull-downs.

III.1.5. Assessment of *in vivo* binding using SILAC worm

To assess the quality of protein-protein interaction screens, it is common practice to validate the results either based on known interactions independently reported in the literature or by using a secondary/orthogonal method to redemonstrate the detected interactions. However, one question that is rarely addressed in reports of protein-protein interactions is whether the identified interaction partners are derived from pre-existing endogenous complexes or whether the bindings were formed after cell lysis. This question is particularly relevant for protein interaction screen using animal tissues, because homogenization breaks down the boundaries of multiple cell and tissue types. As a result, this might lead to bindings formed in solution which would otherwise not be physiologically possible.

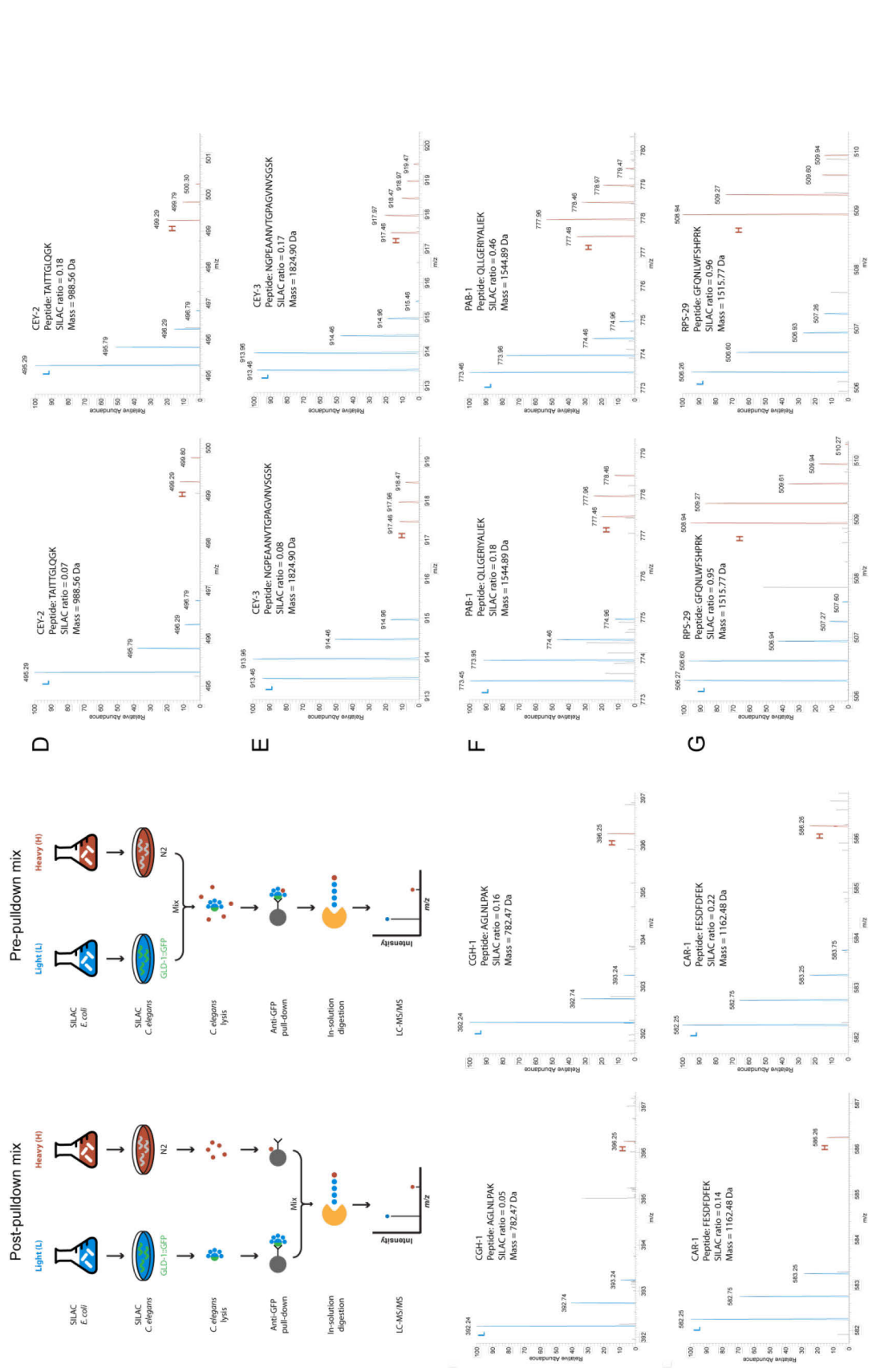

Fig. III.9 | Interaction proteomics detects mainly bindings formed *in vivo*.
(A) Experimental designs using *in vivo* SILAC. Worms harboring GFP fusion proteins were fed on bacteria labeled as "light" (L); control N2 worms were fed on bacteria labeled as "heavy" (H). Samples of the two labeling states were combined either after independent pull-downs with anti-GFP and control beads or prior to the affinity purification steps. (B-F) Representative mass spectra showing peptides of several known interaction partners of GLD-1 (CGH-1, CAR-1, CEY-2, CEY-3, PAB-1). Preferential recovery of the light form in both treatments indicated that minimal binding occurred post-lysis; (G) the non-specific binder, RPS-29, was identified equally in light and heavy forms. A relative increase in the heavy form of PAB-1 when lysates were mixed prior to pull-downs suggested that PAB-1 is in dynamic exchange with a pool of free PAB-1 and/or also binds to GLD-1 *in vitro*. Light peptides are shown in blue; heavy peptides are shown in red.

To address this potential concern, two pull-down experiments were performed using GLD-1::GFP and N2 young adults mixed at different experimental steps (Fig. III.9 A). The GLD-1::GFP worms were metabolically labeled using SILAC as “light” and the wild-type N2 worms as “heavy” (Stoeckius et al., 2014). When the two strains are mixed only at the last step of the experiment, the detected “heavy” proteins will only represent the level of non-specific background bindings to the beads. In contrast, when the two strains are already combined before the lysis step, the detected “heavy” proteins will represent the level of background bindings to the beads plus interactions formed in solution. Therefore, a minimal difference between these two for a detected interaction partner would indicate minimal post-lysis binding, which would strongly suggest that the indentified interaction is derived from pre-existing complexes *in vivo*. As presented in Fig. III.9 B-F, the known interaction partners of GLD-1 (Scheckel et al., 2012) were detected mainly in the “light” form with only minute amounts in the “heavy” form. More importantly, with the exception of PAB-1, the differences for these proteins between the two experimental set-ups are minimal, indicating they are interactions formed *in vivo*. Collectively, these results show that *in vivo* interaction proteomics identifies binding partners derived primarily from pre-existing protein complexes formed *in vivo*.

III.2. Pilot map of *in vivo* protein interactome of *C. elegans* early embryogenesis

III.2.1. Embryo *in vivo* interactome captures known interactions

After establishing the quantitative *in vivo* interaction proteomics method, the present study was extended to investigating the interaction partners of eight key player proteins using established *C. elegans* strains with the following features: 1) the GFP fusion protein either rescues the loss-of-function endogenous counterpart and/or has a localization pattern consistent with the endogenous protein; 2) they are involved in critical developmental processes during early embryogenesis. These bait proteins cover polarity establishment and development (MBK-2, PAR-2, PAR-6, ZEN-4) and P granule assembly/regulation (CAR-1, PGL-3, PIE-1). MBK-1, as a non-essential paralog of MBK-2, was also included.

By quantitative analysis of *in vivo* binding partners of these eight bait proteins in early embryos, a map of 559 interactions among 472 proteins was generated (Fig. III.10; see Fig. VII.1 and Table VII.1 in the supplementary section for an annotated list). This interaction network contains some well-characterized interactions that are important during embryogenesis, including the aforementioned conserved complex of CAR-1/CGH-1 (Boag et al., 2008), the evolutionarily conserved PAR-3/PAR-6/PKC-3 complex (Suzuki and Ohno, 2006), the binding of PAR-6 and LGL-1 (Hoege et al., 2010), the interaction of PAR-2 and PAR-1 (Hoege and Hyman, 2013), the centralspindlin motor complex of ZEN-4 and CYK-4 (Mishima et al., 2002), the association of PGL-3/SEPA-1/EPG-2, which mediates the autophagic degradation of P granules in somatic blastomeres (Tian et al., 2010; Zhang et al., 2009), and the binding of activated MBK-2 to pseudophophatases EGG-3 and EGG-4. Besides previously known bindings, this embryo *in vivo* interactome (IVI) network consists mainly of newly identified interactions that are potential targets for further investigations.

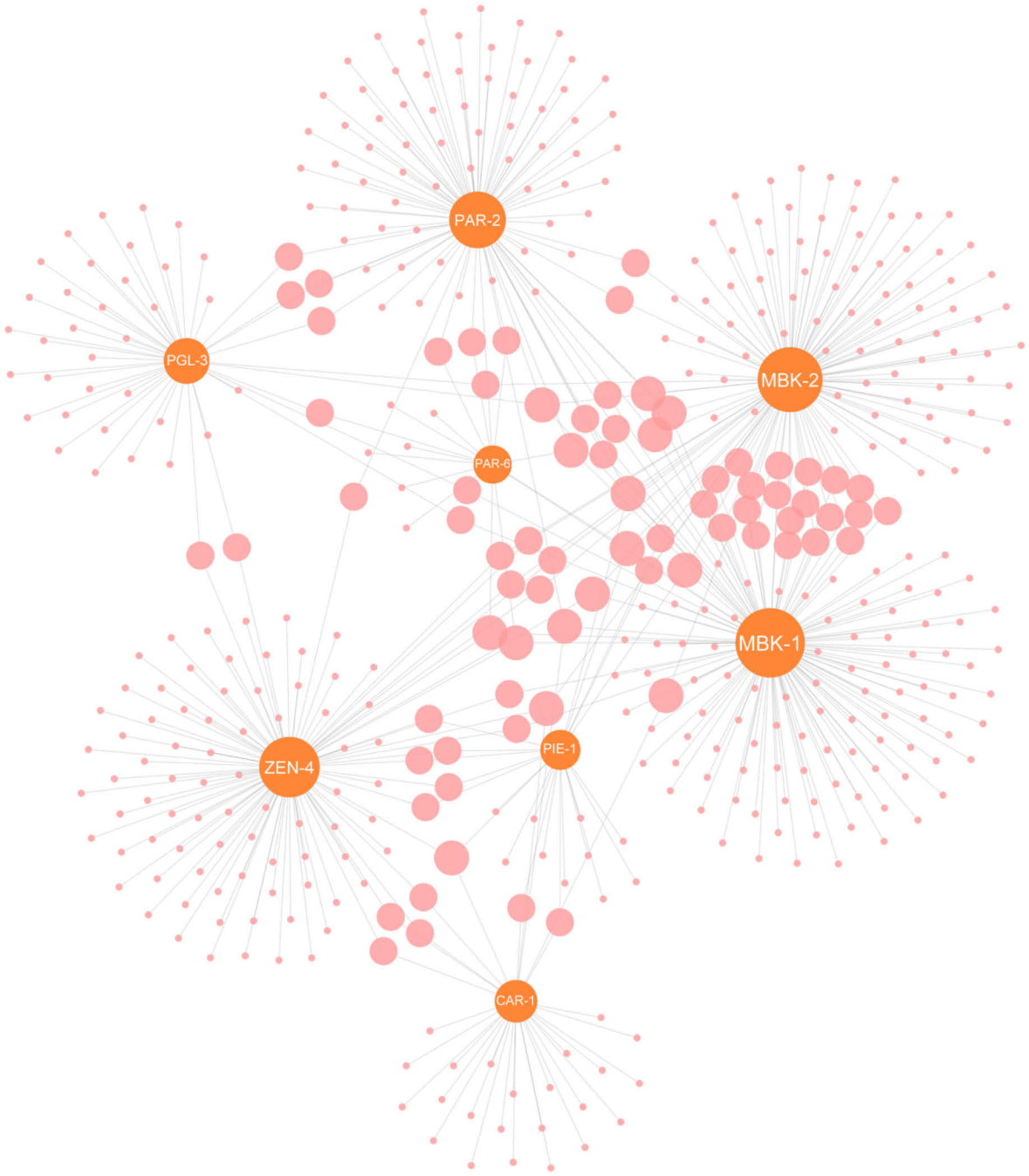

Fig. III.10 | Embryo *in vivo* interactome (IVI) network. Bait (orange) and prey (pink) proteins are shown as nodes sized proportionally to their degree distribution. See Fig. VII.1 and Table VII.1 in the supplementary section for an annotated list.

III.2.2. High biological relevance of embryo *in vivo* interactome

For multi-cellular organisms, systemic mapping of protein-protein interactions by yeast-two hybrid was first performed on *C. elegans* proteins (Li et al., 2004; Walhout et al., 2000). In fact, due to the lack of a cell line system, yeast-two hybrid data remain the primary source of protein-protein interactions reported for *C.*

elegans. The latest release of *C. elegans* protein interactome WI8 (Worm Interactome version 8) consists of interactions assembled from multiple yeast-two hybrid screens (Simonis et al., 2009). Within the same report, a collection of low-throughput interaction data curated from literature (LCI, literature curated interaction) were also presented (Simonis et al., 2009). The LCI dataset represents hand-picked high-quality protein interactions known for *C. elegans*.

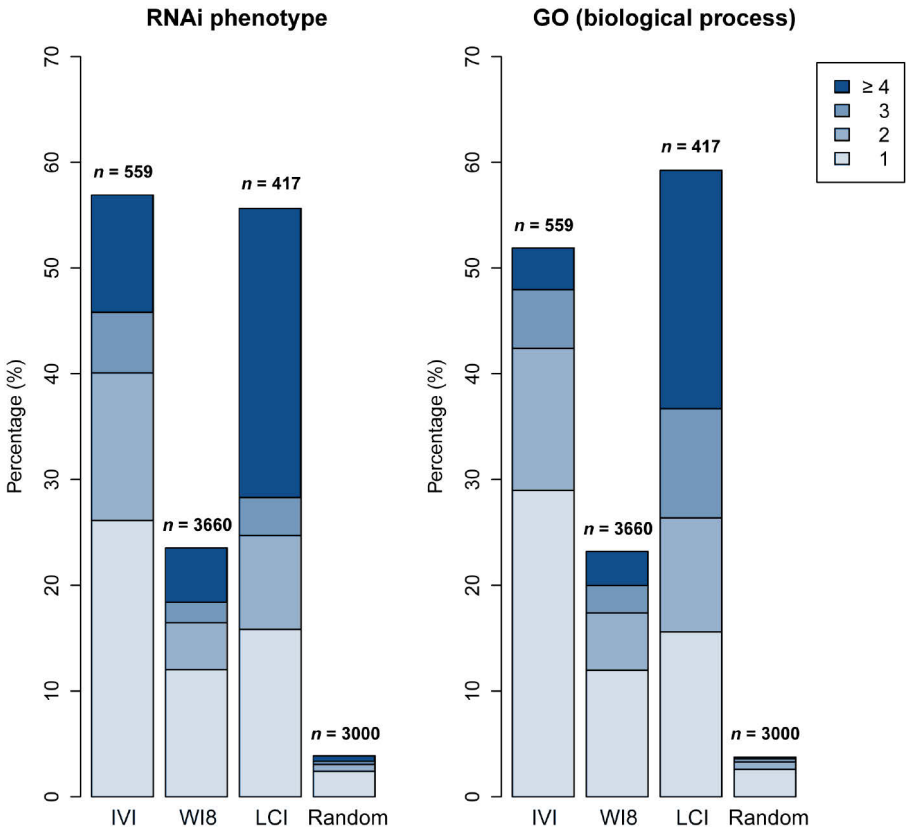

Fig. III.11 | Stacked bar charts showing the percentages of interacting protein pairs that share at least one annotated RNAi phenotype or Biological Process GO term for IVI, WI8 and LCI versus random sets of gene pairs. In the control set, percentages are mean values based on 100 random samples of size 3,000 from all protein-coding genes. The color scale displays the number of shared annotations between each interacting protein pair. IVI, *in vivo* interactome; WI8, Worm Interactome version 8; LCI, literature-curated interaction.

Protein-protein interaction data provide not only physical and spatial information on complex formation but also important hints on functional relationships between binding partners. Since interacting proteins often function in the same biological pathway, abolishing the corresponding genes frequently leads to the same phenotypic effect (Fraser and Plotkin, 2007; Sonnichsen et al., 2005). The large collection of RNAi

phenotype data for *C. elegans*, therefore, provides an invaluable tool to assess the *in vivo* functional associations between genes. To evaluate the functional relevance of the embryo *in vivo* interactome (IVI) generated in the present thesis, the percentage of two interacting proteins that share the same RNAi phenotype(s) was calculated. Indeed, over 50% of all interacting protein pairs from the IVI network share at least one RNAi phenotype, more than 10 times higher than what would be expected by random (< 5%) in the control (Fig. III.11). When compared with the published yeast-two hybrid WI8 and low-throughput LCI datasets, this fraction is on par with the LCI data and considerably higher than the WI8, suggesting the embryo IVI data have equal quality as the independently validated interactions for *C. elegans* (Fig. III.11). This is particularly intriguing, since the assignment of an RNAi phenotype to a gene is purely based on independent experimental observations without prior knowledge on the association between genes. Similar findings were also made for shared gene ontology (GO) terms between interacting protein pairs (Fig. III.11). Altogether, these results indicate the high biological relevance of the embryo *in vivo* interactome data.

III.3. MBK-2 and GEI-12 interaction during early embryogenesis

III.3.1. RNAi screen identified *gei-12* as novel regulator of P granule dynamics

To further explore the value of the embryo *in vivo* interatome dataset, downstream functional experiments were conducted, with a focus on the regulation of P granule dynamics during early embryogenesis. In this context, an RNAi perturbation experiment against the identified interaction partners was set up in order to experimentally uncover new insights into P granule biology.

As already mentioned in the Introduction section of this thesis, the DYRK kinase MBK-2 is one of the critical regulators of oocyte-to-embryo transition leading to the segregation of P granules in the one-cell embryo (Fig. III.12) (Pellettieri et al., 2003). However, in the present thesis, most of the newly identified interaction partners of MBK-2 are not previously known to be involved in P granule formation. Therefore, using GFP fusion PGL-1 and PGL-3 as markers, an RNAi screen against the interaction partners of MBK-2 was performed to search for phenotypes of P granule defects in early embryos. Among a total number of 40 genes individually tested, RNAi knock-down of the gene *gei-12* exhibited a discernible P granule phenotype in the embryos. As displayed in Fig. III.13, while P granules segregated properly with the P lineage in the controls, depletion of *gei-12* by RNAi abolished P granule formation (PGL-1 and PGL-3) during early embryogenesis. In contrast, P granule assembly in the gonadal germline remained unaffected by *gei-12(RNAi)* (Fig. III.13). In the *gei-12(RNAi)* embryos, diffuse GFP signal was found in all embryonic cells, while in the germline blastomere only minute amounts of discernible granules were

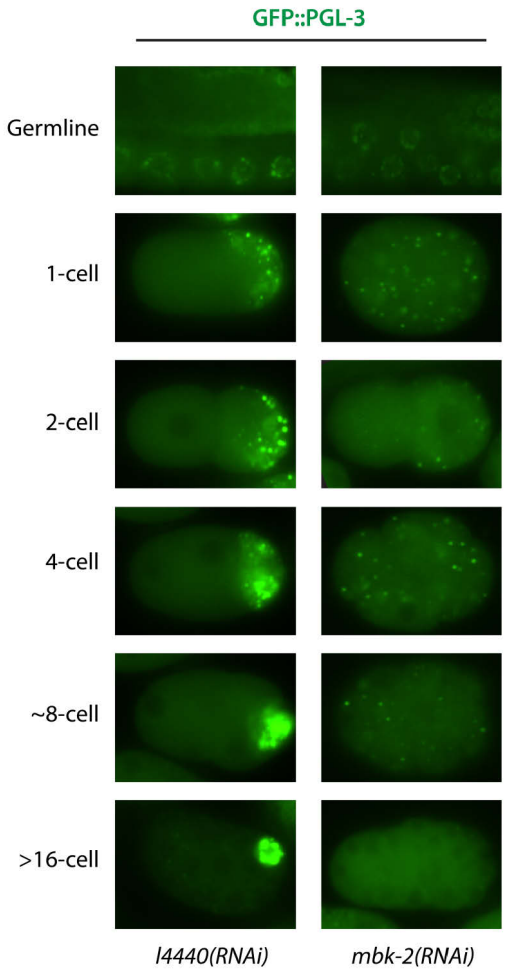

Fig. III.12 | Depletion of *mbk-2* by RNAi disrupts P granule disassembly during early embryogenesis.. *L4440(RNAi)* served as control.

observed which disappeared with embryonic cell divisions (Fig. III.13). These observations were also confirmed by staining against endogenous P granule proteins using K76 antibodies (Fig. III.14).

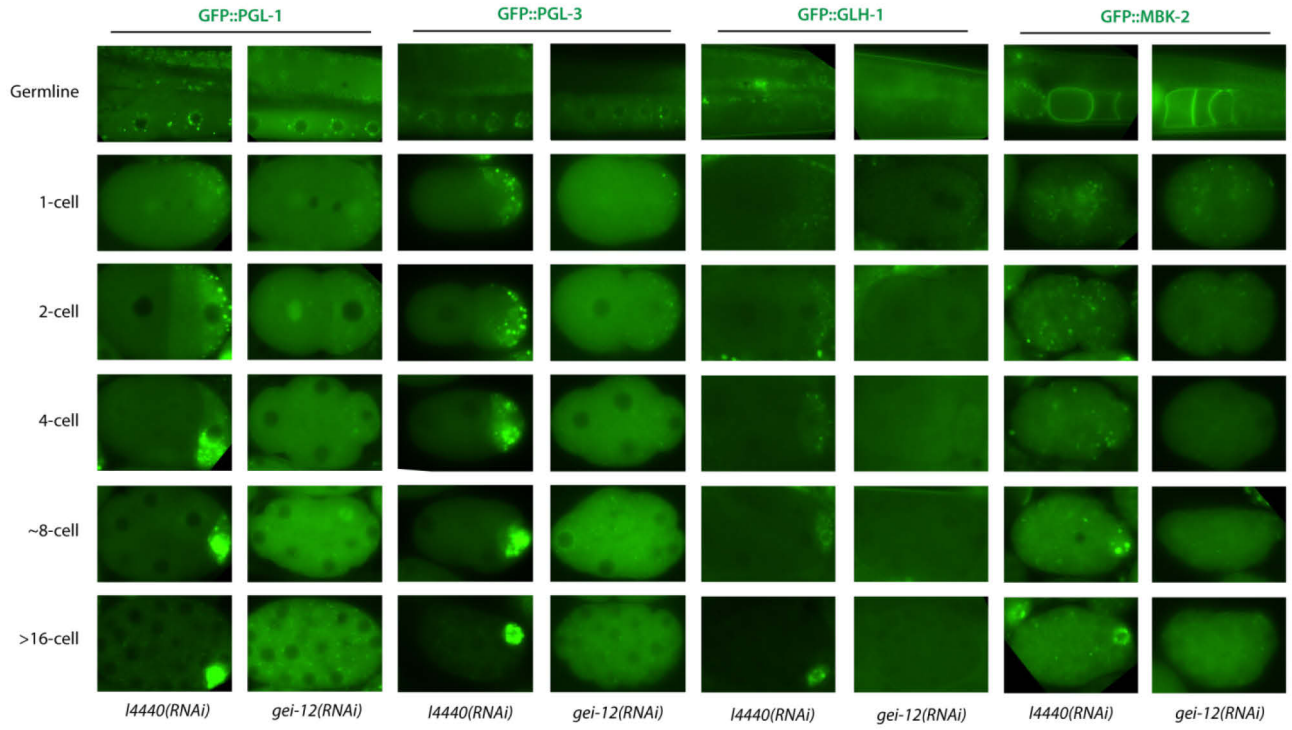

Fig. III.13 | GEI-12 is a novel protein required for P granule assembly. Depletion of *gei-12* by RNAi affected P granule assembly specifically during early embryogenesis but not prior to fertilization, as visualized in live embryos harboring GFP fusions of several proteins that localize to P granules. *L4440(RNAi)* served as control.

To further assess the impact of *gei-12* depletion on P granules, another constitutive component of P granules, GLH-1, was chosen as a marker. According to the hierarchy of P granule assembly based on genetic studies (Kawasaki et al., 1998; Updike and Strome, 2010), *glh-1* acts upstream of *pgl-1* and *pgl-3* in the assembly chain. Upon depletion of *gei-12*, a similar assembly defect phenotype was observed for the fusion protein GFP::GLH-1 (Fig. III.13). This result suggests that during early embryogenesis *gei-12* is upstream of *glh-1* and *pgl-1/3* in the formation of P granules. Interestingly, the localization of MBK-2 to the germline blastomere was also sabotaged by *gei-12(RNAi)*, indicating *gei-12* is also required for the recruitment of MBK-2 to the P granules in early embryos (Fig. III.13). However, further experiments showed that knock-down of *gei-12* by RNAi did not affect the segregation of a few other posteriorly localized proteins (CAR-1, PIE-1, POS-1 and SPN-4) to the P lineage (Fig. III.15). In summary, these results suggest that *gei-12* is indispensible for the germline segregation of a subgroup of P granule proteins, including PGL-1/3, GLH-1 and

MBK-2, specifically during early embryogenesis, and that *gei-12* is not crucial for the asymmetric distribution of other P granule components.

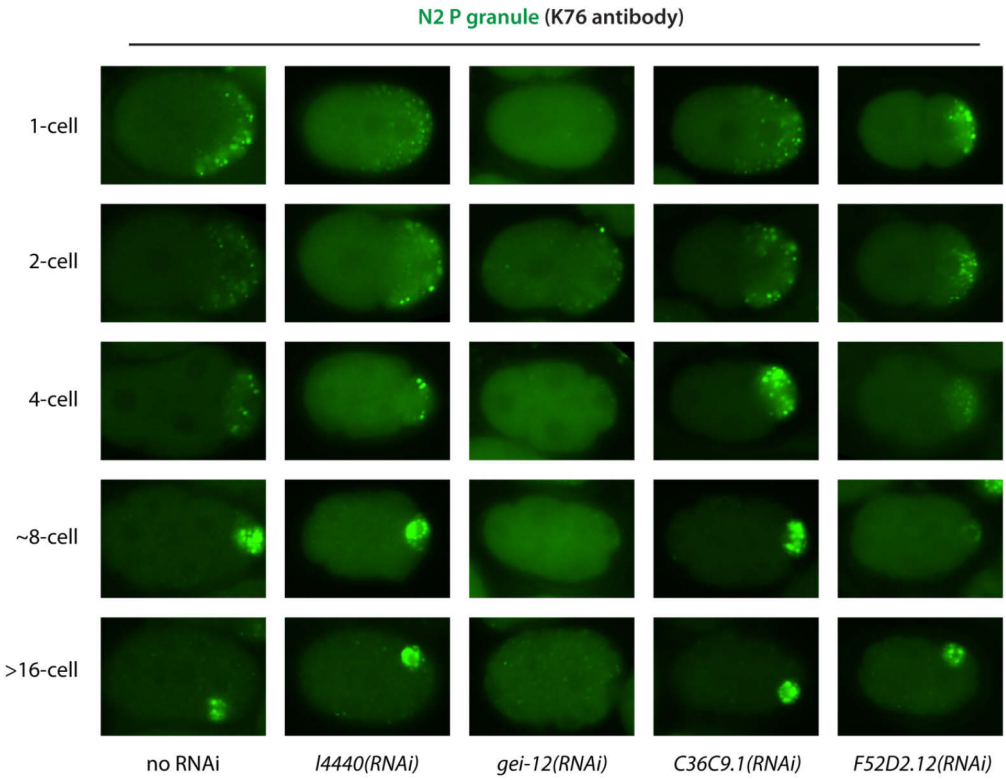

Fig. III.14 | Depletion of *gei-12* by RNAi disrupts P granule formation in N2 early embryos. Knocking down the paralogs of *gei-12* individually by RNAi in N2 embryos did not impair P granule assembly. Fixed embryos were stained with K76 antibody against P granules.

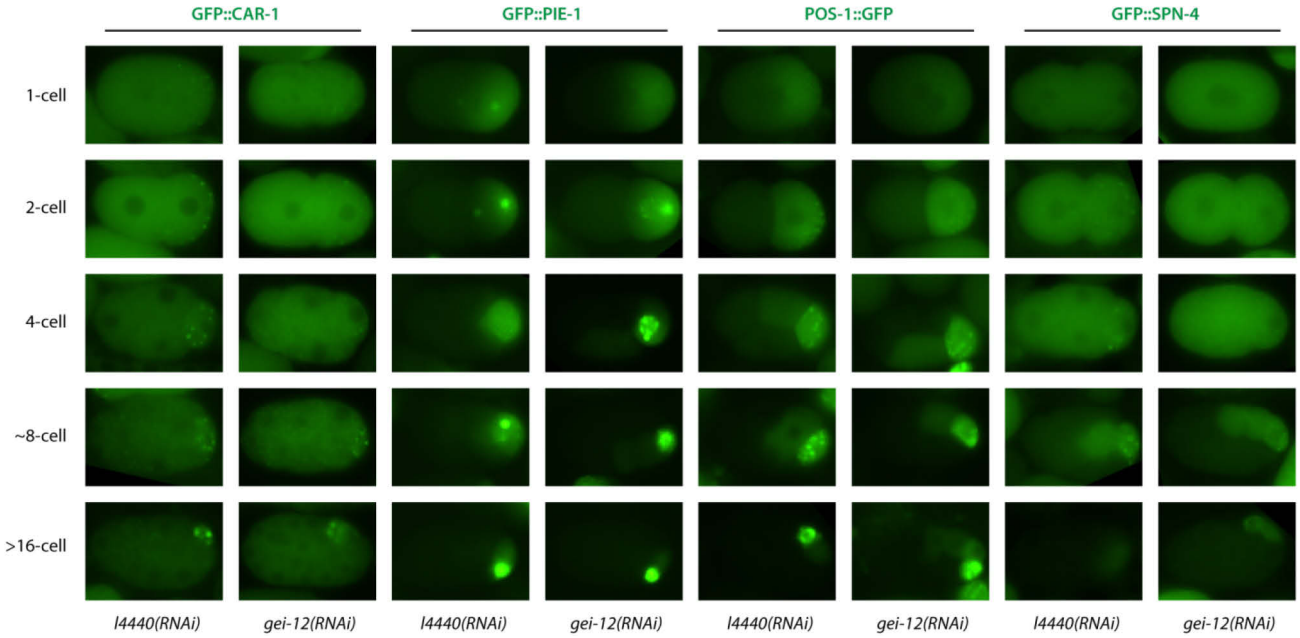

Fig. III.15 | Depletion of *gei-12* by RNAi does not affect asymmetric localization of some other known P granule components (CAR-1, PIE-1, POS-1 and SPN-4) during early embryogenesis.

III.3.2. *Gei-12* and its paralogs

Despite appearing to be a *C. elegans* specific gene with no clear orthologs even within nematoda, *gei-12* has two known paralogs (C36C9.1 and F52D2.12), all of which are located on the X chromosome with a nearly identical exon-intron structure and C36C9.1 being closer to *gei-12* regarding sequence similarity (Fig. III.16). In order to study the potential roles of these paralogs in P granule formation, mutants with large deletions (presumably loss-of-function null mutants) were used to investigate their functional association with *gei-12*.

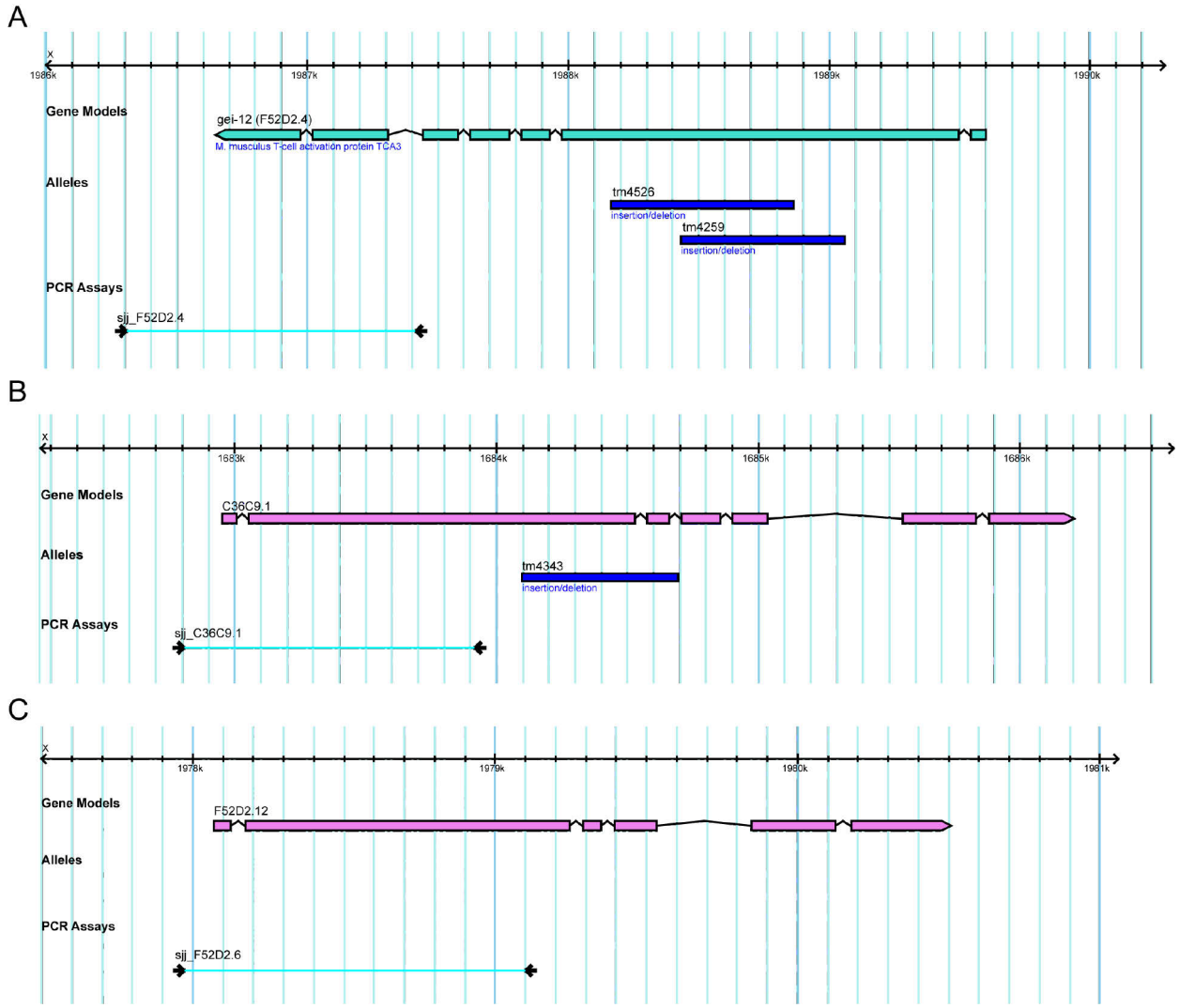

Fig. III.16 | Gene models of *gei-12* and its paralogs C36C9.1 and F52D2.12. Deleted regions of mutant alleles (*tm4526*, *tm4259*, *tm4343*) are shown in blue. Regions targeted by individual RNAi clones are shown in the "PCR Assays" category. Retrieved and reproduced from Wormbase.

Among these deletion mutants, *gei-12(tm4526)* leads to a large in-frame deletion while both *gei-12(tm4259)* and *C36C9.1(tm4343)* cause an additional frame-shift (Fig. III.16). Interestingly, by using K76

antibody staining, P granules in embryos appeared to be unaffected in individual deletion mutants (Fig. III.17). In addition, depleting the paralogs C36C9.1 and F52D2.12 individually by RNAi in wild-type N2 worms also did not lead to a P granule assembly defect (Fig. III.14). Together with the observation of a P granule phenotype in *gei-12(RNAi)* embryos (Fig. III.14), these results suggest that there is a certain degree of functional redundancy among GEI-12 and its paralogs and that the *gei-12* RNAi clone (sjj_F52D2.4) can efficiently knock down multiple paralogs. Importantly, the sequence regions that can be potentially targeted by this particular RNAi clone are indeed highly similar among *gei-12* and its paralogs. Moreover, when combined with RNAi knock-down of C36C9.1, both *gei-12* mutants displayed a P granule defect with a higher penetrance than RNAi depletion of *gei-12* in wild-type background.

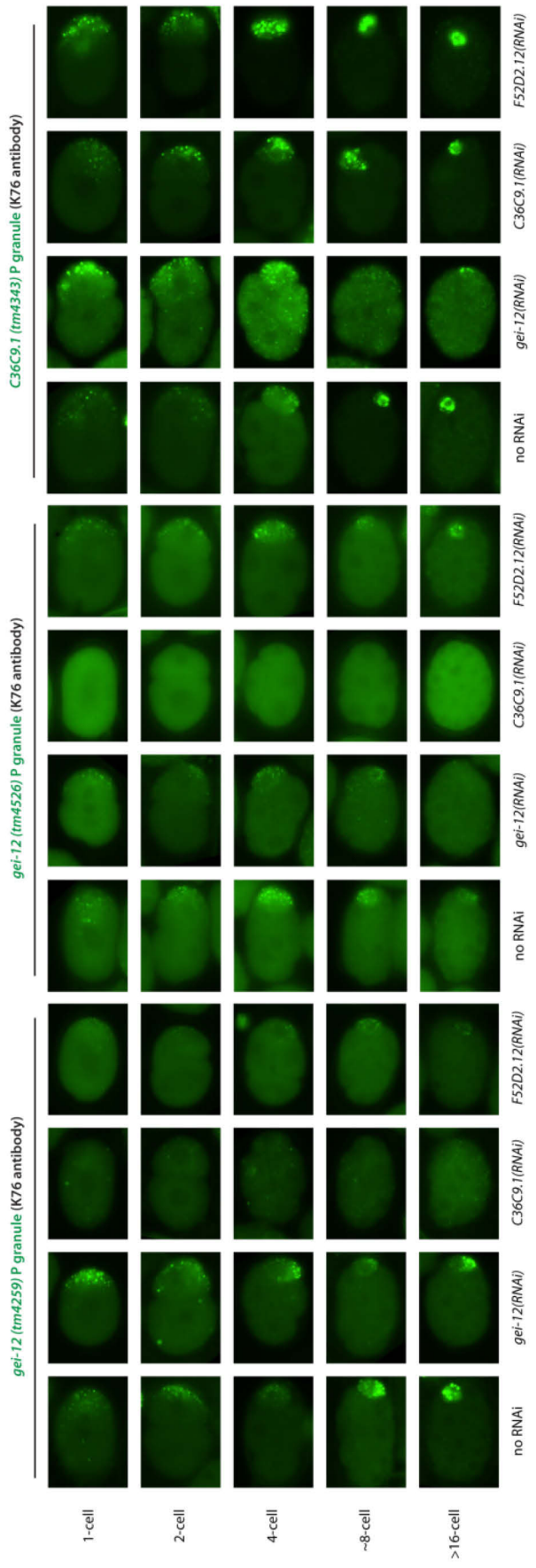

Fig. III.17 | GEI-12 and its paralogs play partially redundant roles in P granule assembly. Strains harboring mutations in *gei-12*(*tm4259* or *tm4526*) and its paralog *C36C9.1*(*tm4343*) were fixed and stained with K76 antibody against P granules, without or with RNAi against each member of the *gei-12* family. The P granule phenotype elicited by *gei-12(RNAi)* was considerably weaker in the two *gei-12* mutant backgrounds than in wild type (Fig. III.14). RNAi against either *gei-12* paralog affected the P granule phenotype of *gei-12* mutants to variable extents, but had no obvious effect in the *tm4343* mutant background or in wild-type (Fig. III.14).

III.3.3. GEI-12 segregates with the germline in *C. elegans*

To further understand the physiologic functions and localization of GEI-12 in *C. elegans*, a GFP fusion GEI-12 knock-in strain was generated. Consistent with the aforementioned roles on P granule assembly, the distribution of GEI-12 throughout development matches the typical localization pattern of a P granule component. As depicted in Fig. III.18, GEI-12 localizes to developing germ cells with a characteristic peri-nuclear pattern and segregates with the P lineage during embryogenesis. Thus, GEI-12 is a maternally-loaded germline factor which regulates P granule assembly during embryonic development.

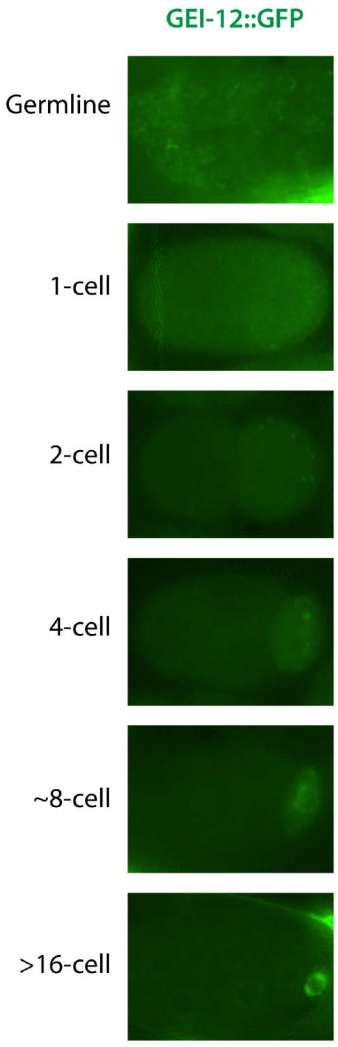

Fig. III.18 | GEI-12 displays a typical P granule localization pattern in *C. elegans*, segregating with the germline and the P lineage throughout development.

III.3.4. Evidence for role of GEI-12 phosphorylation in P granule dynamics

The data presented above identified GEI-12 as a critical protein required for P granule assembly. However, the mechanism how GEI-12 is involved in this process is unclear. The protein appears to be worm-specific with no clear orthologs in other species. Due to this lack of sequence homology and annotated domains, it is difficult to speculate about its functions. A striking feature of the GEI-12 sequence is the high percentage (19.4%) of serine/threonine residues (Fig. III.19). The fact that GEI-12 was identified to be an MBK-2 kinase interaction partner makes it tempting to speculate that MBK-2 phosphorylates GEI-12 on some of these serine/thronine residues. For DYRK2/3 kinases, the mammalian homologs of MBK-2, the phosphorylation consensus sequences ($RX_{2-3}S/TP$) have been well characterized by *in vitro* experiments (Campbell and Proud, 2002). Replacement of the N-terminal arginine ($RX_{2-3}S/TP$) by a lysine residue ($KX_{2-3}S/TP$) abolishes the phosphorylation activity by DYRK2/3 (Campbell and Proud, 2002), although such a replacement by a similar basic amino acid residue is tolerated for phosphorylation site

recognition by some basophilic serine/threonine protein kinases (Pinna and Ruzzene, 1996). When this N-terminal arginine restriction is loosened to K/RX$_{2\text{-}3}$S/TP, GEI-12 indeed contains one putative motif (residues: KATKS(582)P) matched to the DYRK2/3 phosphorylation consensus sequences, suggesting that GEI-12 could be a potential substrate of MBK-2 kinase.

Additionally, the P granule assembly defect phenotype of *gei-12(RNAi)* is reminiscent of the PP2A phosphatase regulatory subunit *pptr-1* mutant (Gallo et al., 2010), suggesting that GEI-12 could be a downstream target of PP2A and potentially as its substrate for dephosphorylation. Further lines of evidence supporting these speculations include: 1) PPTR-2, another PP2A regulatory subunit and paralog of PPTR-1, was identified as interaction partner of MBK-2 in the present study; 2) PPTR-2 was previously found to interact with GEI-12 by yeast-two hybrid (Simonis et al., 2009).

Fig. III.19 | GEI-12 protein sequence analysis. Serine/threonine residues (19.4%) shown in green; predicted low complexity regions highlighted in magenta. Prediction was performed using the SEG algorithm with default parameters (Wootton and Federhen, 1996).

Collectively, during early embryogenesis, MBK-2 potentially phosphorylates GEI-12 which leads to disassembly of P granule in the anterior half of the embryo, whereas PP2A potentially dephosphorylates GEI-12 and consequently promotes P granule assembly in the posterior. If this extrapolation were true, depleting *mbk-2* might be able to rescue the P granule assembly defect of the *pptr-1* mutant. To test if this is really the case, RNAi was used to knock down *mbk-2* in the *pptr-1* defective mutant. Indeed, loss of *mbk-2* could restore P granule formation in *pptr-1* mutant embryos (Fig. III.20), indicating a counteracting relation between *mbk-2* and *pptr-1* over P granule assembly during early embryogenesis. It needs to be mentioned

that there remains a lack of direct evidence in the present thesis to demonstrate that GEI-12 is phosphorylated by MBK-2 and dephosphorylated by PP2A *in vivo*, due to unavailability of antibody against GEI-12 and the *gei-12::gfp* transgene being quickly silenced for expression *in vivo*. Unfortunately, extensive efforts to generate a stably integrated GEI-12::GFP worm line without silencing for further biochemistry experiments were not successful.

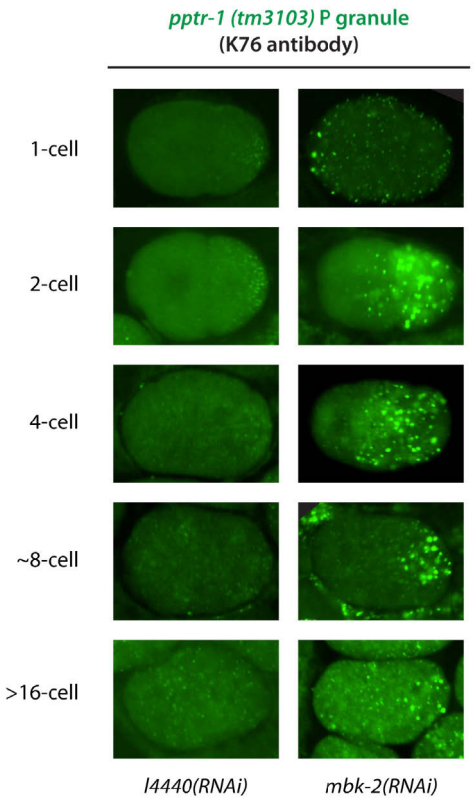

Fig. III.20 | Depletion of *mbk-2* by RNAi restores P granule formation in *pptr-1* mutant embryos. P granules do not segregate normally to the posterior due to loss of cell polarity upon *mbk-2* depletion. *L4440(RNAi)* served as control. Experiment performed by Patricia Cipriani, New York University.

III.3.5. GEI-12 is required for germline maintenance

Sterility assay performed by Patricia Cipriani, New York University.

Simultaneous depletion of multiple P granule components has been shown to result in sterility after more than one generation of RNAi treatment (Updike et al., 2014). Likewise, when L1 larvae were treated with *gei-12* RNAi, they remained fertile upon reaching adulthood; however, further application of *gei-12* RNAi to the progeny lead to about 5% sterility in the F1 generation and ~20% in the F2 generation at 15°C, and these percentages rose to ~25% and ~35% when the assay was carried out at 25°C (Fig. III.21 A). In the hermaphrodites, the sterile worms showed hollow uteri and limited germline proliferation with no gametes (also in males; Fig. III.21 B and C). In summary, GEI-12 is a key germline factor that is also needed for the proliferation of germline, and this requirement becomes greater with temperature. A previous study reached a similar conclusion for the P granule components MEG-1 and MEG-2 (Leacock and Reinke, 2008).

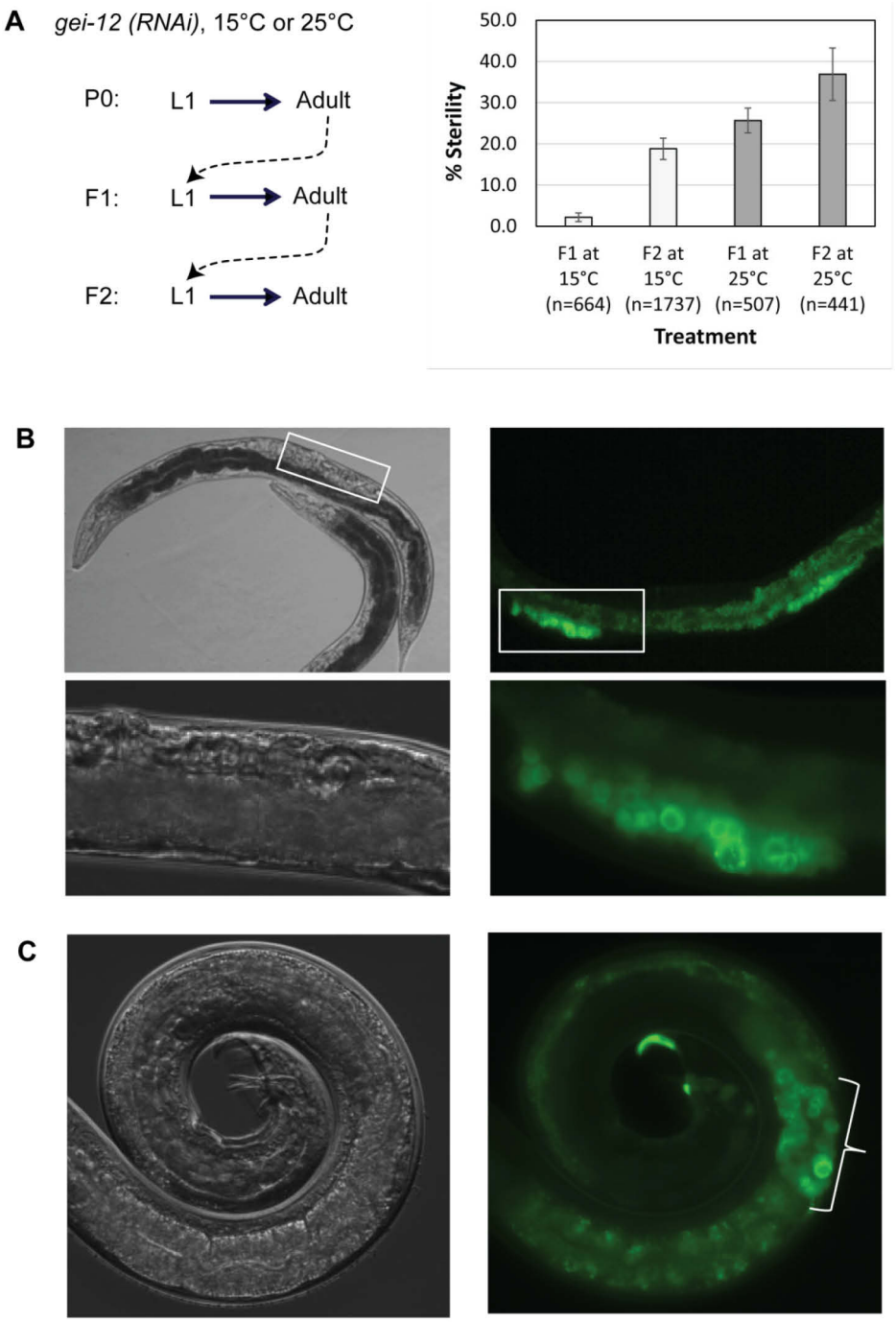

Fig. III.21 | GEI-12 is required for germline proliferation and gamete formation. (A) Protocol for continuous generational RNAi of *gei-12* by feeding. Average sterility rates among adult progeny in first (F1) and second (F2) filial generations, with standard error. Adults from the P0 generation were not sterile; sterility among adult progeny increased with generation and temperature. (B-C) Examples of adult hermaphrodite (B) and male (C) progeny after generational *gei-12(RNAi)* treatment. Light micrographs (left) and GFP::PGL-1 (right) show reduced germlines with fewer PGL-1 positive cells in both sexes. Sterile hermaphrodites also show empty uteri devoid of embryos. Bottom panels in (B) show magnified view of highlighted boxes in upper panels. In (C), bracket spans region of defective male germline.

III.4. GEI-12 granule formation in mammalian cells

III.4.1. Analysis of full-length GEI-12 protein

The findings of GEI-12 and its relationships with MBK-2 and PP2A pose an interesting question on whether phosphorylation/dephosphorylation is a general mechanism regulating RNP granule formation. Based on sequence feature analysis, GEI-12 contains no previously known functional domains other than several low complexity regions (Fig. III.19). Also, by using the IUPred algorithm (Dosztanyi et al., 2005), GEI-12 is predicted to be highly disordered along most parts of the entire protein, especially in its N-terminal region (Fig. III.22).

Many disordered proteins are believed to be able to adopt various conformations upon binding to proteins with a stable structure (Meszaros et al., 2009). According to the ANCHOR algorithm (Meszaros et al., 2009), which predicts the regions of a protein capable of a transformation from being disordered to ordered upon binding to a globular domain, most parts of the disordered sequences of GEI-12 are predicted to be binding regions to structured proteins (Fig. III.23). This prediction result suggests that GEI-12 could potentially be a flexible scaffold protein that can bind to multiple proteins.

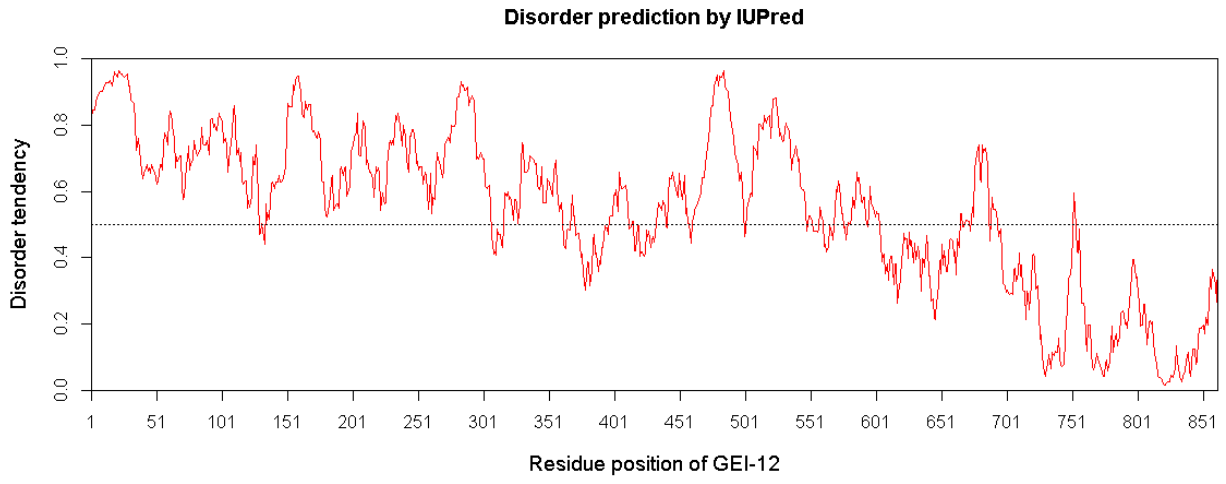

Fig. III.22 | Disorder prediction of GEI-12 by IUPred algorithm. Prediction type: "long disorder". Dotted line indicates the threshold value of 0.5, above which suggests disordered/unstructured regions.

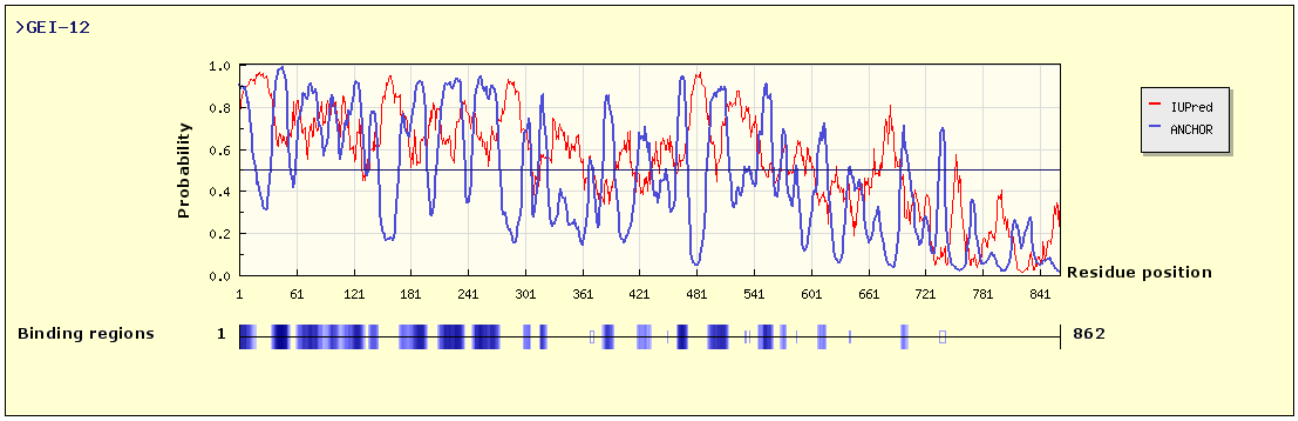

Fig. III.23 | Binding region prediction within the disordered parts of GEI-12 by ANCHOR algorithm. The upper panel: probabilities calculated by ANCHOR (blue) are superimposed on the IUPred disorder probabilities (red); the threshold probability value is 0.5. The lower panel: long blue vertical bars indicate binding regions; short vertical bars/rectangles are regions filtered out by the algorithm.

When expressed in HEK293T cells fused to an N-terminal EGFP, GEI-12 formed granular structures of various sizes in the cytosol (Fig. III.24 A). This is intriguing since there are no known homologs of GEI-12 in mammals. To further understand these granules, a SILAC-based pull-down experiment against EGFP::GEI-12 was performed to identify the proteins binding to GEI-12 (Fig. III.24 B). Among its ~300 specific interaction partners, over two-thirds turned out to be known RNA-binding proteins, including FXR-1/2, Staufen, translation initiation factors and ribosomal proteins, suggesting that these GEI-12 enriched granular structures in mammalian cells are RNP granules. This is consistent with the role of GEI-12 in *C. elegans* as expected from the composition of P granules. Furthermore, these proteins (~10%) were also significantly enriched for previously known components of various types of RNP granules (p value = 6×10^{-21}, hypergeometric test) (Kato et al., 2012). One particularly interesting discovery among these interaction partners is the PP2A phosphatase subunits, including the catalytic subunit (PPP2CA), structural subunits (PPP2R1A, PPP2R1B) and, specifically, several regulatory subunit B56 subfamily members (PPP2R5A, PPP2R5C, PPP2R5E). Most of the proteins co-purifying with GEI-12 were not affected by nuclease treatment, suggesting that they are not RNA dependent (Fig. III.24 C).

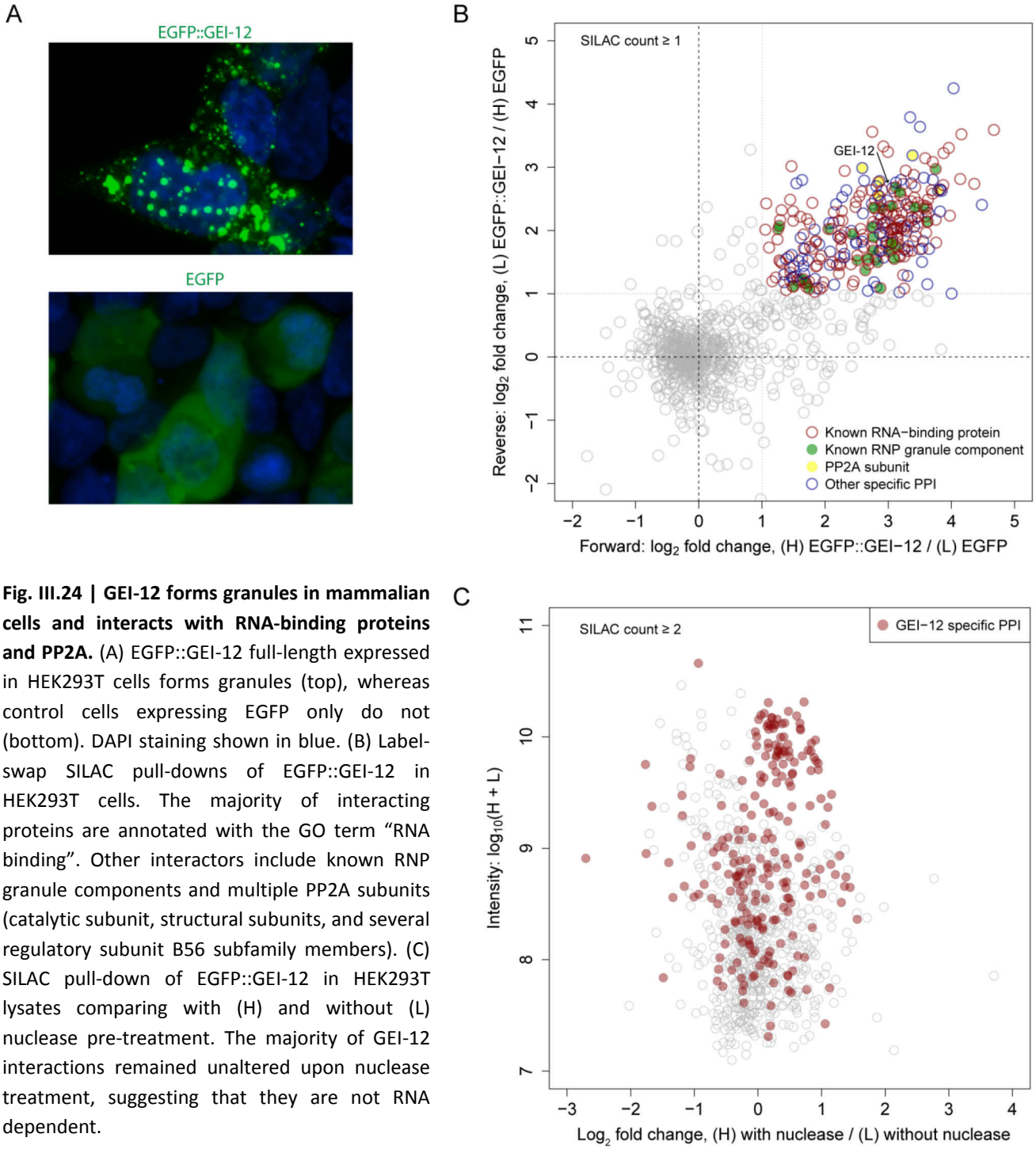

Fig. III.24 | GEI-12 forms granules in mammalian cells and interacts with RNA-binding proteins and PP2A. (A) EGFP::GEI-12 full-length expressed in HEK293T cells forms granules (top), whereas control cells expressing EGFP only do not (bottom). DAPI staining shown in blue. (B) Label-swap SILAC pull-downs of EGFP::GEI-12 in HEK293T cells. The majority of interacting proteins are annotated with the GO term "RNA binding". Other interactors include known RNP granule components and multiple PP2A subunits (catalytic subunit, structural subunits, and several regulatory subunit B56 subfamily members). (C) SILAC pull-down of EGFP::GEI-12 in HEK293T lysates comparing with (H) and without (L) nuclease pre-treatment. The majority of GEI-12 interactions remained unaltered upon nuclease treatment, suggesting that they are not RNA dependent.

III.4.2. Behavior of GEI-12 fragments in mammalian cells

To further understand the biochemical properties of GEI-12, different fragments of the full-length protein were cloned into an expression vector containing an N-terminal EGFP. When expressed in HEK293T cells, these fragments displayed rather different localization and distribution patterns (Fig. III.25). All fragments containing the N1 region (residues 2-143) were exclusively or, for the N-terminal fragment, partially localized to the nucleus. Indeed, GEI-12 contains two predicted nuclear localization signal motifs at residues 15-20 (RRKRGG) and 291-296 (RKKSRG) (Dinkel et al., 2014), which might in part explain the nuclear localization patterns. The N-terminal fragment also formed granules in the cytosol which were generally bigger than those of the full-length proteins. In contrast, the C-terminal fragment appeared to be solely cytosolic and homogeneously distributed.

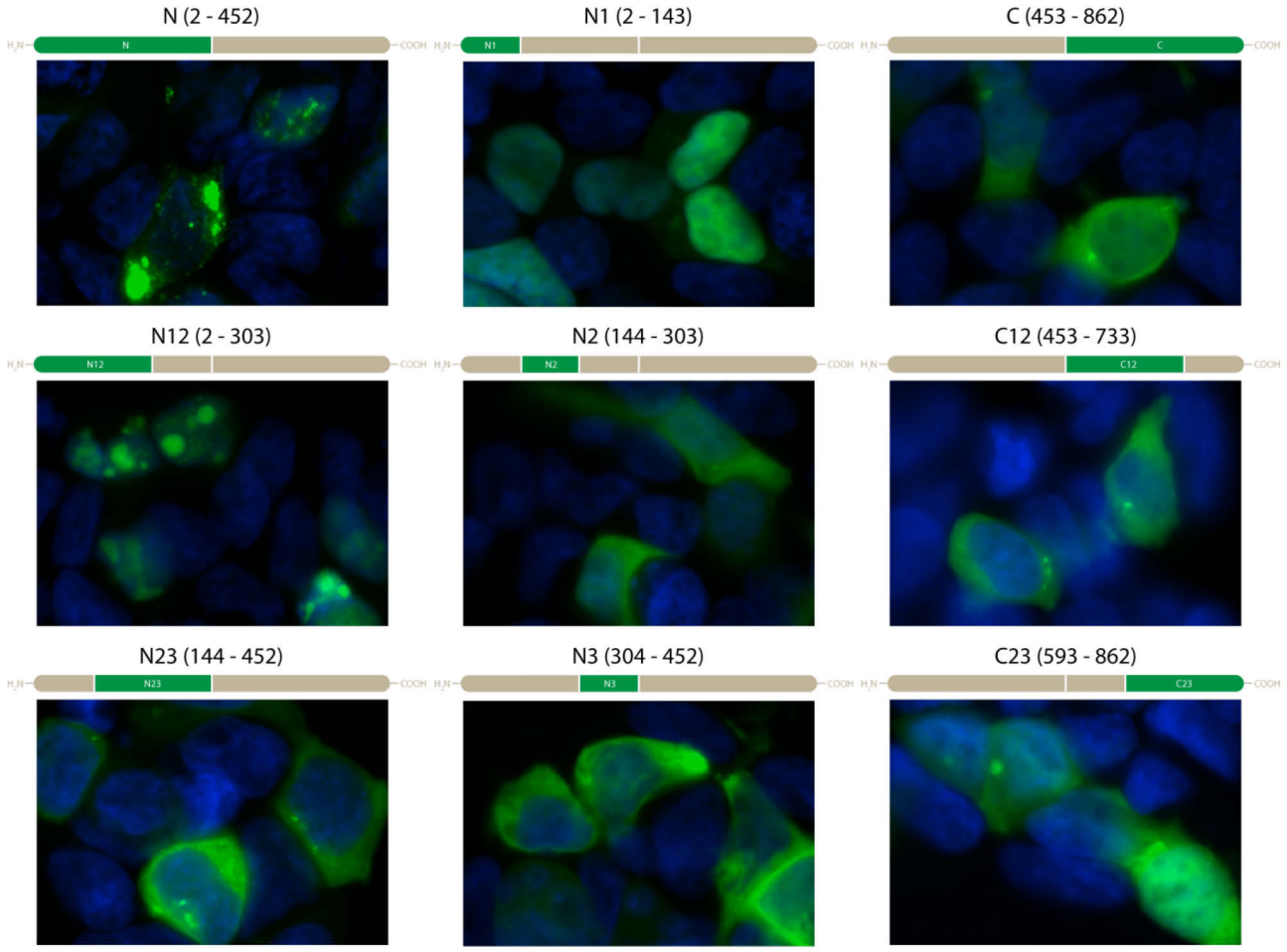

Fig. III.25 | Expression of GEI-12 fragments fused to an N-terminal EGFP in HEK293T cells. Amino acid residue positions of GEI-12 fragments are indicated in parentheses (begin – end).

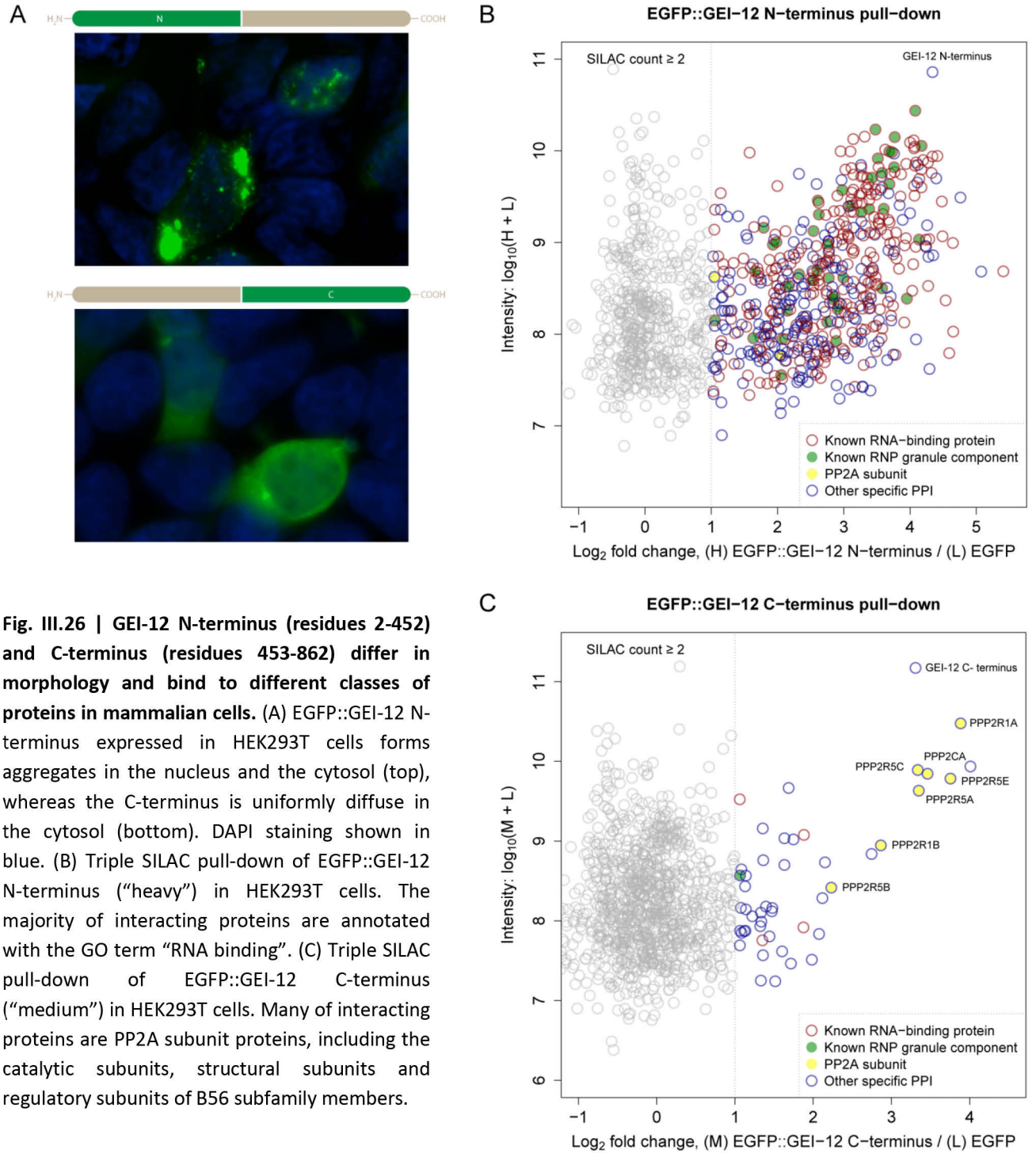

Fig. III.26 | GEI-12 N-terminus (residues 2-452) and C-terminus (residues 453-862) differ in morphology and bind to different classes of proteins in mammalian cells. (A) EGFP::GEI-12 N-terminus expressed in HEK293T cells forms aggregates in the nucleus and the cytosol (top), whereas the C-terminus is uniformly diffuse in the cytosol (bottom). DAPI staining shown in blue. (B) Triple SILAC pull-down of EGFP::GEI-12 N-terminus ("heavy") in HEK293T cells. The majority of interacting proteins are annotated with the GO term "RNA binding". (C) Triple SILAC pull-down of EGFP::GEI-12 C-terminus ("medium") in HEK293T cells. Many of interacting proteins are PP2A subunit proteins, including the catalytic subunits, structural subunits and regulatory subunits of B56 subfamily members.

Given the very different distribution patterns of the N- and C-terminal parts of GEI-12 expressed in mammalian cells, one interesting question is: do they also interact with different types of proteins? Indeed, SILAC pull-down experiments against the N-terminus of GEI-12 revealed that the majority of its interaction partners were RNA-binding proteins (Fig. III.26 B), suggesting that the N-terminus is required for RNP granule formation in the presence of other RNA-binding proteins. Conversely, the C-terminal part of GEI-12 appeared

to bind mainly the PP2A phosphatase subunits including the catalytic & structural subunits and the regulatory subunit B56 subfamily members (PPP2R5A, PPP2R5B, PPP2R5C, PPP2R5E), indicating that the C-terminus of GEI-12 is the docking site for PP2A and/or is the region of GEI-12 regulated by PP2A dephosphorylation activity (Fig. III.26 C). Since the disordered regions of GEI-12 are mainly distributed in the N-terminal half of GEI-12 (Fig. III.22), whether this is one reason behind the differences in their distribution pattern as well as binding partners is an interesting question for further exploration.

IV Discussion

IV.1. *In vivo* interaction proteomics

Quantitative *in vivo* interaction proteomics is an emerging field that allows detection of specific protein interactions with high confidence in an organism under a physiologic/near-physiologic context (Andlauer et al., 2014; Hanack et al., 2015; Lundby et al., 2014). The simplicity of a label-free approach retains a high degree of quantification accuracy while being simpler and cheaper compared to the sometimes prohibitive cost of *in vivo* SILAC labeling of animals (Cox et al., 2014; Hubner et al., 2010; Hubner and Mann, 2011). The versatility is enormous, and such an *in vivo* approach can in principle be extended to studying interactions in pathologic conditions, aging, organogenesis, as well as the dynamical changes of interactions during these conditions, and so on. In the present thesis, a label-free quantitative *in vivo* interaction proteomic approach was developed for *C. elegans* embryos. With further optimization, it should be possible to apply the method to investigating protein interactions during other developmental stages and to follow the changes of these interactions.

The present method makes use of GFP fusion *C. elegans* strains which are fortunately already widely available for a large collection of proteins. Moreover, a visible tag offers further utilities in downstream functional and validation assays. In principle, this method can be applied to any GFP fusion strain. However, one should also be aware of the potential limitations of the method. Although it has been demonstrated that known endogenous interaction partners can be efficiently captured using the current method with high sensitivity, not all the GFP fusion strains available in the research community are fully representative of the endogenous counterparts. This could be due to multiple integrations of the transgene, functional deficits of the fusion protein especially in non-rescued strains, non-endogenous expression control and so on. To partly solve these problems, a recent large collection of GFP-tagged fosmids that allow for endogenous-level expression of the fusion protein under *cis* regulatory control could be very useful (Sarov et al., 2012). Another solution could be to take advantage of the genome editing technique using the CRISPR-Cas9 system for seamless insertion of an affinity tag next to the endogenous gene location (Hsu et al., 2014; Tzur et al.,

2013). Such strains should be able to closely mimic the behavior of the endogenous proteins at a physiologic level. Nevertheless, an affinity tag as well as its tagging position (N- or C-terminal) could interfere with or change the binding preference of the bait protein. Potential artifacts are therefore embedded in the data. With regard to the throughput of the method, it is still rather low, and thus extending it efficiently to a "genome-wide" scale of bait proteins remains difficult.

The present thesis also tried to determine if the detected interaction partners were derived from pre-existing complexes *in vivo* or from bindings formed post lysis in solution. Using SILAC labeled transgenic and wild-type worms plus mixing at different experimental steps, it was demonstrated that the identified interactors for one test strain were indeed derived primarily from protein complexes formed *in vivo*. This was an important step to biochemically assess the quality of the pull-down data. However, it could not be ruled out that for other fusion proteins there could still be detected interactions formed post lysis. Interestingly, this apparently important issue is rarely addressed in published interaction studies. In fact, the idea has been used previously to assess the binding dynamics of protein complexes (Mousson et al., 2008). If cell lysates of different SILAC states are mixed before affinity purification, the same binding protein of different SILAC labels might swap in between the bait proteins. In contrast, combining independent pull-downs of different SILAC states only at the final step minimizes any potential swap between labels. A comparison of these two approaches could reveal the dynamic behaviors of some interactions. In the present thesis, looking from a slightly different viewpoint, the very same technique can be applied to investigating the extent of post-lysis effect on complex formation in pull-down experiments.

Although single-step purification and a highly efficient single-chain anti-GFP antibody have been employed in the current method, complex integrity may not be completely preserved upon tissue disruption. Therefore, a crosslinking step might be helpful to rescue protein complex components that would otherwise become dissociated after lysis (Subbotin and Chait, 2014). Furthermore, an interesting technical addition to a crosslinking approach could be to utilize photoactivatable amino acids that allow photo-crosslinking *in vivo* and potentially also in living animals (Suchanek et al., 2005). Of note, a recent report demonstrated fast production of highly efficient engineered camelid antibodies (Fridy et al., 2014). In particular, their chimeric

single-chain antibodies have been shown capable of capturing 90% of the target protein complex in just 10 minutes. The reduction in sample preparation time would help to preserve weak or transient interactors after affinity purification. Recent development in tissue-specific proteomics using non-canonical amino acids also offers the possibility to pinpoint one step further the interactions in a specific tissue or cell type of *C. elegans* (Yuet et al., 2015).

IV.2. Embryo *in vivo* interactome network

In the present thesis, a pilot *C. elegans* embryo interactome map was created for 8 bait proteins. Despite the relatively small number of bait proteins, a sizable network consisting of 559 interactions among 472 proteins was generated. Importantly, this network captures well-characterized interactions known for their critical roles during embryonic development. Likewise, comparison with yeast-two hybrid and literature-curated interactions indicates high biological relevance of the *in vivo* interactome data. Although this is not a fully automated high-throughput approach as the robotically operated yeast-two hybrid technique, it is not unrealistic to expand the screen to a much larger number of bait proteins. Such a bigger network could present finer details on the cooperativity of proteins and the modular organization of protein machines during embryogenesis.

Recent development in high-throughput techniques has led to a previously unthinkable magnitude of expansion of big data. Although a global or systems view is increasingly appreciated, the overwhelming complexity of information from large-scale datasets is often times a daunting challenge to tackle. While systems biologists continue to innovate to extract useful information at a non-traditional level, how to move on from large-scale screen data to classical functional studies to explore new biology remains a highly tricky task. In *C. elegans*, the high-throughput nature of RNAi perturbation screen offers an *in vivo* solution for a rapid secondary assay for functional follow-ups. Another solution is through computational integration of information at multiple levels to achieve, ironically, data reduction (Stroedicke et al., 2015). In the present thesis, only potential new regulators of P granules have been investigated in the functional assays. A

comparison with some unpublished genetic interaction data yields several highly interesting candidates for experimental follow-ups involving other biological processes. The high functional relevance of the current dataset plus the demonstrated usefulness in finding new regulator of P granule dynamics during embryogenesis indicate that the embryo *in vivo* interaction data could be a rich resource for the *C. elegans* research community for further exploration.

IV.3. P granule dynamics during *C. elegans* embryogenesis

RNP granules are membraneless compartments contributing to post-transcriptional regulation. Cytoplasmic RNP granules such as stress granules and processing bodies are involved in RNA metabolism under various cellular conditions, thereby modulating gene expression programs according to functional needs of the cell (Anderson and Kedersha, 2006; Kedersha et al., 2013). In a similar fashion, components of the developmentally regulated P granules in *C. elegans* have been shown to be involved in gene expression control related to the development of the germline lineage (Mello et al., 1996; Tabara et al., 1999; Voronina et al., 2011). P granules are believed to be sites of sequestration of certain maternal RNA, regulating their stability and expression in response to dynamic developmental cues. Knowing which RNA species undergo such expression controls is therefore of great interest to developmental biologists. Perhaps even more interesting questions are: how are different RNA species selected to be incorporated into these granules? What are the responsible proteins that determine the outcome of this selection? Although not much is currently known, one could speculate that certain transcript sequence features, structural motifs or RNA modifications might assist in the recognition process (Pagano et al., 2007).

At the preceding step, the molecular mechanisms governing P granule formation remain elusive. During embryogenesis, P granule assembly has been shown to follow a genetic hierarchy: *deps-1* -> *glh-1* -> *pgl-1/3* -> *ife-1* (Updike and Strome, 2010). However, this simple pathway does not capture all the complex regulatory mechanisms that have been found involved in the control of P granule dynamics. The anterior-posterior gradient of RNA-binding protein MEX-5 in the one-cell embryo appears closest to be the direct

regulator that controls the asymmetric partition of P granules (Nishi et al., 2008). The current model indicates that the anteriorly partitioned MEX-5 promotes dissolution of P granules in the anterior half of the embryo and that these dissolved P granule proteins consequently accumulate in the posterior half and coalesce into bigger granules (Brangwynne et al., 2009). Interestingly, no study has so far demonstrated how MEX-5 biochemically dissolves P granules. With concrete evidence from all the previously published perturbation experiments, it is highly unlikely that the gradient of MEX-5 and dissolution of P granules are pure coincidence. Indeed, it has been speculated that MEX-5 might be the carrier/scaffold for another effector protein, such as the polo kinase PLK-1 or PLK-2, which undertakes the direct task of P granule dissolution. Other effector proteins might also be involved. The precise mechanism requires further investigations.

The DYRK kinase MBK-2, regulator of oocyte-to-embryo transition, has been shown to be required for asymmetric segregation of P granules during early embryogenesis in the absence of a clear spindle/polarity defect, suggesting that MBK-2 mediates anterior P granule dissolution (Pang et al., 2004; Pellettieri et al., 2003). In contrast, the PP2A phosphatase subunit PPTR-1 is essential for P granule assembly during mitosis in the one-cell embryo (Gallo et al., 2010). In the present thesis, it was demonstrated that depletion of *mbk-2* by RNAi rescued the loss of P granule phenotype in the *pptr-1* null mutant, suggesting a genetic interaction between *mbk-2* and *pptr-1* in controlling P granule assembly/disassembly during early embryogenesis.

The protein GEI-12, newly identified in the present thesis, segregates with the P lineage, displays a P granule localization pattern and is an *in vivo* binding partner of MBK-2. Unfortunately, generating a stably integrated *C. elegans* line of GEI-12::GFP fusion was unsuccessful despite extensive efforts. Therefore, an *in vivo* binding assay between GEI-12 and PPTR-1 could not be performed in *C. elegans* embryos. However, pull-down experiments against GEI-12 expressed in mammalian cells clearly showed the interaction between GEI-12 and PP2A subunits, including in particular the B56 regulatory subunit subfamily, which are the homologs of PPTR-1. Given also the finding that *gei-12* appears to be located upstream of *glh-1* and *pgl-1/3* in the P granule assembly pathway, all these results suggest that P granule disassembly and assembly specifically during early embryogenesis are in part controlled by phosphorylation (MBK-2) and

dephosphorylation (PPTR-1), respectively, at the potential targets of GEI-12 and its paralogs. In summary, a hypothetical model is proposed as presented in Fig. IV.1: in the one-cell stage embryo after fertilization, MBK-2 phosphorylates GEI-12 and as a result leads to disassembly of P granules in the anterior pole of the embryo, whereas in the posterior GEI-12 is dephosphorylated by PP2A and consequently promotes P granule assembly. This hypothetical model was independently supported by a very recent report showing that MBK-2 can phosphorylate recombinant GEI-12 *in vitro* and that the phosphorylation status of GEI-12 fusion protein *in vivo* is influenced by *mbk-2* and *pptr-1* (Wang et al., 2014). Of note, whether MBK-2 and PPTR-1 are the direct effector proteins on the phosphoyrlation status of GEI-12 *in vivo* is yet to be experimentally confirmed. In this context, it is likely that more kinases and phosphatases are involved. Finding the *in vivo* interaction partners of GEI-12 in embryos using the method developed in the present thesis may uncover new candidates.

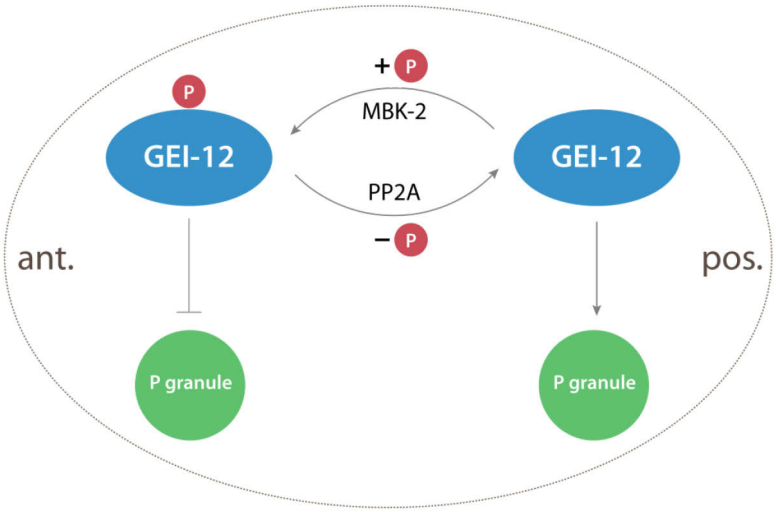

Fig. IV.1 | Hypothetical model of P granule disassembly and assembly controlled via GEI-12 by phosphorylation (MBK-2) in the anterior and dephosphorylation (PP2A) in the posterior of the one-cell embryo.

Previous model of P granule formation is based on the discovery that PGL proteins (PGL-1 and PGL-3) are able to autonomously induce RNP granule formation both in mammalian cells and ectopically in *C. elegans* somatic cells in the absence of other germline-specific factors, suggesting that PGL proteins can self-associate and serve as scaffolds for recruitment of other P granule components (Hanazawa et al., 2011). Findings in the present thesis support a new model of phosphorylation control of P granule development

during early embryogenesis (Fig. IV.1). In this model, GEI-12 acts as the seed protein for initiating P granule assembly and the phosphorylation status determines its aggregation property and, in turn, the condensation/dissolution of P granules. GEI-12 is a worm-specific protein with no homologs in mammals, based on sequence conservation assessment. However, in mammalian cells in the absence of PGL proteins, over-expressed GEI-12 can form granules which are enriched for RNA-binding proteins and known RNP granule components. This suggests that despite the lack of apparent conservation at the sequence level the fundamental biophysical properties required for granule formation are conserved across species during evolution. Of note, instead of granule induction, it cannot be excluded that GEI-12 localizes to existing RNP granules in mammalian cells. However, given the relatively large size of some of the GEI-12 granules in the cytosol (Fig. III.24 A), it is unlikely they were already formed in the absence of GEI-12. A microscopy experiment comparing the localization of GEI-12 interaction partners in mammalian cells before and after GEI-12 expression would partially address this question.

Recent studies have repeatedly emphasized the emerging role of low complexity regions in RNP granule/hydrogel formation (Kato et al., 2012; Kwon et al., 2013; Wippich et al., 2013). Low complexity regions are composed of residues of low compositional complexity ranging from homotypic repeats of single amino acids to mosaic clusters of only a few amino acid types (Wootton, 1994). Due to their low content of residue information, these regions are by default filtered out in alignment algorithms, such as BLAST. *In vitro* studies have demonstrated that low complexity regions alone are sufficient to form hydrogels and are able to retain other RNA-binding proteins (Kato et al., 2012). While this provides evidence to understand the basic mechanism of RNP granule formation, the regulatory mechanisms *in vivo* are likely to be far more complex. For instance, in the embryos of *lgg-1* mutant where autophagosomal degradation of PGL proteins is impaired, ectopic PGL protein granules appeared in the somatic cells from 16-cell stage onwards (Zhang et al., 2009). However, inactivation of *sepa-1* in these mutant embryos leads to diffuse distribution of the ectopic PGL proteins in the somatic cells (Zhang et al., 2009). These results suggest that despite the ability of some granule proteins to self-assemble into granule/hydrogel there appear to be a cohort of promoting and

suppressing factors of granule formation, whose complex interplay regulates the assembly and disassembly of P granules *in vivo*.

IV.4. RNP granule assembly and disassembly

The current thesis presents evidence to support a hypothetical model in which the phosphorylation status of GEI-12 orchestrates the stability of P granules, thereby regulating their segregation during embryogenesis. Dephosphorylated form of GEI-12 appears to favor P granule assembly, based on the previous report that PP2A phosphatase regulatory subunit PPTR-1 is required for P granule assembly in the one-cell embryo (Gallo et al., 2010) and current finding that GEI-12 form granules and binds to PP2A subunit proteins in mammalian cells. Furthermore, when expressed in mammalian cells, the C-terminal half of GEI-12, the less disordered part of the protein, has a uniformly diffuse distribution in the cytosol but appears to be responsible for the binding to PP2A phosphatase subunits, whereas the N-terminal half forms aggregates of mainly RNA-binding proteins. With these lines of evidence plus the unusually high fraction of serine/threonine residues (19.4%) in the GEI-12 sequence, it is tempting to speculate that dephosphorylated GEI-12 nucleates granule formation whose stability is maintained by PP2A phosphatases. Likewise, upon phosphorylation by MBK-2 or other kinases which changes the net charge/charge distribution of GEI-12, granule structure is destabilized and dissolves as a result.

Indeed, unphosphorylated GEI-12 appears to be primarily positively charged, which is complementary to the negatively charged RNA molecules (Fig. IV.2 A). In a thought experiment, introducing the negatively charged phosphate groups to GEI-12 in the extreme case could completely change its charge distribution and turn it into a highly negatively charged protein (Fig. IV.2 B). This could then presumably disrupt the electrostatic interactions between GEI-12 and RNA, leading to granule dissolution. Given the high number of potential phosphorylation sites, it is a challenging task to use classical site-specific mutagenesis to pinpoint which one or combination of serine/threonine residues of GEI-12 are critical for granule formation.

Alternatively, combining live cell imaging with perturbation experiments using selective phosphatase inhibitors might help to understand how PP2A stabilizes these RNP granules.

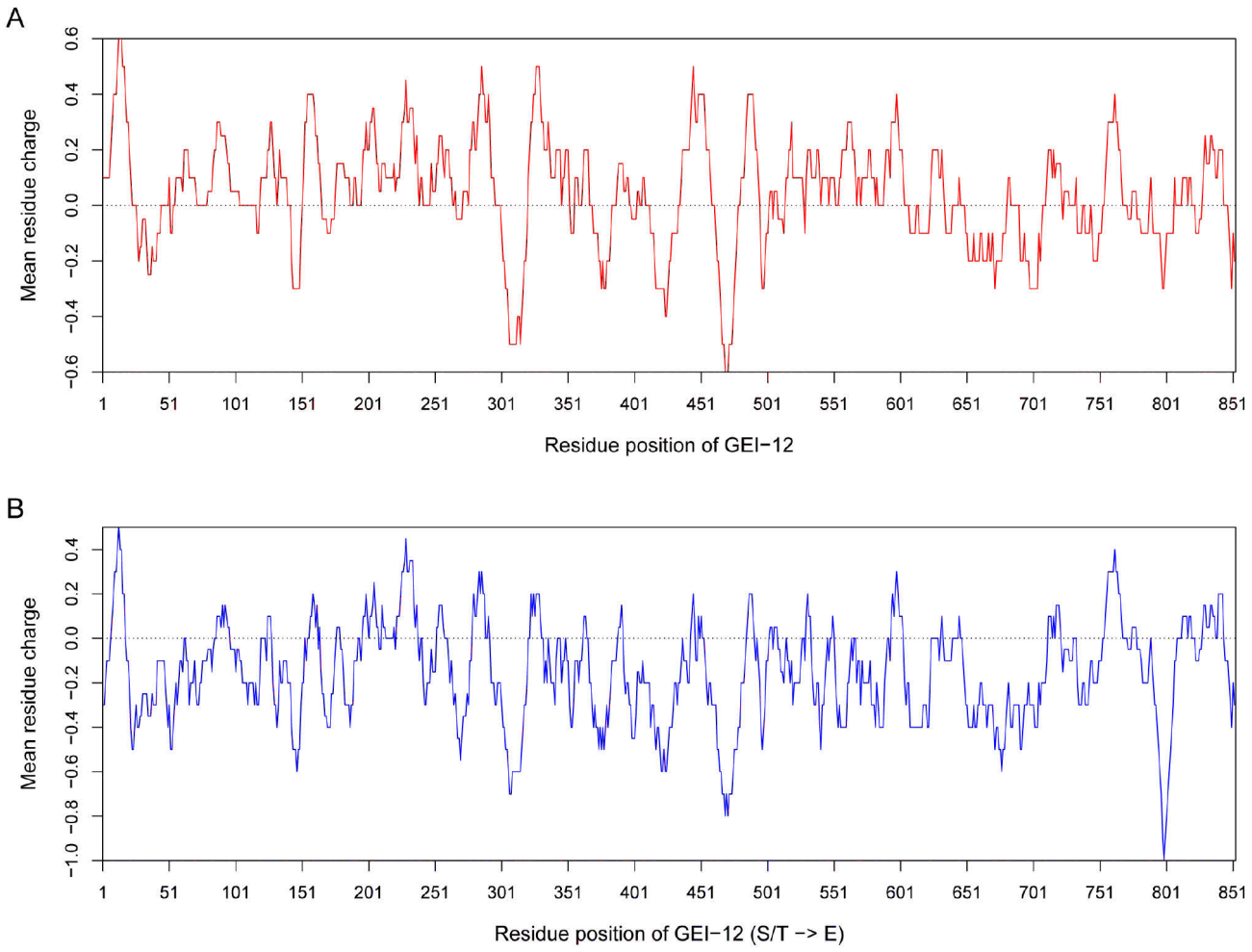

Fig. IV.2 | Charge distribution of wild-type GEI-12 (A) and imaginary "highly phosphorylated" version of GEI-12 (B) where all serine/threonine residues have been mutated to glutamic acid to mimic phosphorylation. Displayed are the mean charges of a sliding window of 10 amino acids along the length of the protein.

RNP granules are known to resemble physical gels where the polymeric matrix is held by weak non-covalent forces (Hyman et al., 2014). This suggests that a scaffold of interlinked structure with a certain degree of orderliness might be present in the granules, whose stability could be affected by factors such as temperature, ionic strength and post-translational modifications (Kwon et al., 2013; Nott et al., 2015). Due to the important role of phosphorylation in regulating signal transduction, recent studies have highlighted the relationship between phosphorylation and hydrogel/granule formation (Kwon et al., 2013; Wippich et al., 2013). This is a significant step for understanding how cells regulate the dynamics of membrane-free

compartmentalization. Wippich and co-authors found that the kinase activity of DYRK3 is required for cytoplasmic stress granule dissolution. Another study demonstrated that the self-polymerized low complexity region of the RNA-binding protein FUS functioned as transcriptional activation domain, and that the FUS hydrogel could trap the disordered C-terminal region of RNA polymerase II large subunit in its unphosphorylated form but not when it was phosphorylated (Kwon et al., 2013). These results led to their proposed mechanistic model how phosphorylation regulates the recruitment and release of RNA polymerase II to and from actively transcribed genes (Kwon et al., 2013). In summary, these previous findings are consistent with the currently proposed model of P granule dynamics controlled by kinase and phosphatase interplay.

Finally, there remain many interesting questions regarding regulation of P granule dynamics. To name a few, P granules are liquid-like droplets that can drip and fuse together (Brangwynne et al., 2009). However, the emulsion of many small P granules in the cytosol raises an apparent question: what factors prevent these granules of close proximity from fusing to become one big aggregate, which is in theory thermodynamically favorable, similar to the process of Ostwald ripening (Brangwynne, 2013). Another one is: despite no clear gross morphological change, what is the in and out of various transient P granule components contributing to the functional role of P granules during development. Answers to many of these questions could reveal further details of RNP granule biology which will provide not only the underlying knowledge on normal cell physiology but also increasingly the medical implications for understanding and treating human diseases (Kedersha et al., 2013; Li et al., 2013b; Simpson-Holley et al., 2011; Somasekharan et al., 2015; Speese et al., 2012).

V Conclusions and Outlook

Studying protein-protein interactions is instrumental in elucidating complex biological processes. In the post-genomic era, comprehensive mapping of protein-protein interactions provides critical information on the mechanical and functional organization of the proteome. Although large-scale interaction networks are already available for most model organisms, studying protein-protein interactions *in vivo* can produce arguably the most biologically relevant data. In the present thesis, a new affinity purification approach using label-free quantitative interaction proteomics has been developed for investigating *in vivo* protein-protein interactions in *C. elegans* embryos. Using this method, the first pilot network of *in vivo* interactions of *C. elegans* embryogenesis was generated. Importantly, this network captures well-characterized known complexes and new interactions of high biological relevance. From a technical perspective, one interesting extension of the method is to study the dynamics of *in vivo* protein interactions over a developmental timeline. From a biological perspective, this dataset could be a rich resource for future biological follow-up studies. Also, expanding the network to a bigger number of bait proteins can deliver more elaborated details on the molecular and functional co-operation that regulates early *C. elegans* development.

Follow-up studies of the embryo interaction data by RNAi revealed a novel protein, GEI-12, which is an essential regulator of P granule dynamics specifically during early embryogenesis. Further functional characterization of GEI-12 led to the hypothetical model that MBK-2 kinase and PP2A phosphatase (PPTR-1) regulate the phosphorylation status of GEI-12 (and its homologs) and consequently orchestrate the formation of P granule during embryogenesis. However, more kinases and phosphatases may be involved in controlling GEI-12 phosphorylation. Finding out these potential players will help elucidate this complex process.

Although GEI-12 appears to be worm-specific with no clear homologs in other species, the ability of GEI-12 to form RNP granules in mammalian cells suggests the fundamental biophysical properties of RNP granule formation are conserved over evolution. Therefore, it would be interesting to see if proteins possessing similar sequence and functional characteristics of GEI-12 exist also in mammals. The unphosphorylated form

of GEI-12 appears to be favorable for granule formation whose stability seems to be maintained by PP2A-mediated activities. It would be interesting to reconstitute granule assembly using recombinant GEI-12 of various phosphorylation states and assess their abilities to form hydrogels in solution as well as the kinetics of bindings to other P granule components. In this context, perhaps what may be even more interesting is to develop new strategies to modulate RNP granule formation, which might have implications for novel therapeutic options in clinical practice. In summary, the fundamental mechanisms governing RNP granule assembly and disassembly have just begun to be understood. More research efforts are needed to uncover these universal principles that underlie many of the basic biological phenomena.

VI References

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198-207.

Ahrne, E., Molzahn, L., Glatter, T., and Schmidt, A. (2013). Critical assessment of proteome-wide label-free absolute abundance estimation strategies. Proteomics *13*, 2567-2578.

Anderson, P., and Kedersha, N. (2006). RNA granules. The Journal of cell biology *172*, 803-808.

Andlauer, T.F., Scholz-Kornehl, S., Tian, R., Kirchner, M., Babikir, H.A., Depner, H., Loll, B., Quentin, C., Gupta, V.K., Holt, M.G.*, et al.* (2014). Drep-2 is a novel synaptic protein important for learning and memory. eLife *3*.

Angrand, P.O., Segura, I., Volkel, P., Ghidelli, S., Terry, R., Brajenovic, M., Vintersten, K., Klein, R., Superti-Furga, G., Drewes, G.*, et al.* (2006). Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling. Molecular & cellular proteomics : MCP *5*, 2211-2227.

Aranda, S., Laguna, A., and de la Luna, S. (2011). DYRK family of protein kinases: evolutionary relationships, biochemical properties, and functional roles. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *25*, 449-462.

Audhya, A., Hyndman, F., McLeod, I.X., Maddox, A.S., Yates, J.R., 3rd, Desai, A., and Oegema, K. (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in Caenorhabditis elegans. The Journal of cell biology *171*, 267-279.

Bartoi, T., Rigbolt, K.T., Du, D., Kohr, G., Blagoev, B., and Kornau, H.C. (2010). GABAB receptor constituents revealed by tandem affinity purification from transgenic mice. The Journal of biological chemistry *285*, 20625-20633.

Boag, P.R., Atalay, A., Robida, S., Reinke, V., and Blackwell, T.K. (2008). Protection of specific maternal messenger RNAs by the P body protein CGH-1 (Dhh1/RCK) during Caenorhabditis elegans oogenesis. The Journal of cell biology *182*, 543-557.

Boag, P.R., Nakamura, A., and Blackwell, T.K. (2005). A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in C. elegans. Development *132*, 4975-4986.

Boxem, M., Maliga, Z., Klitgord, N., Li, N., Lemmens, I., Mana, M., de Lichtervelde, L., Mul, J.D., van de Peut, D., Devos, M.*, et al.* (2008). A protein domain-based interactome network for C. elegans early embryogenesis. Cell *134*, 534-545.

Brangwynne, C.P. (2013). Phase transitions and size scaling of membrane-less organelles. The Journal of cell biology *203*, 875-881.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proceedings of the National Academy of Sciences of the United States of America *108*, 4334-4339.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Campbell, L.E., and Proud, C.G. (2002). Differing substrate specificities of members of the DYRK family of arginine-directed protein kinases. FEBS letters *510*, 31-36.

Cheeks, R.J., Canman, J.C., Gabriel, W.N., Meyer, N., Strome, S., and Goldstein, B. (2004). C. elegans PAR proteins function by mobilizing and stabilizing asymmetrically localized protein complexes. Current biology : CB *14*, 851-862.

Cheeseman, I.M., Niessen, S., Anderson, S., Hyndman, F., Yates, J.R., 3rd, Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. Genes & development *18*, 2255-2268.

Cheng, K.C., Klancer, R., Singson, A., and Seydoux, G. (2009). Regulation of MBK-2/DYRK by CDK-1 and the pseudophosphatases EGG-4 and EGG-5 during the oocyte-to-embryo transition. Cell *139*, 560-572.

Coordinators, N.R. (2015). Database resources of the National Center for Biotechnology Information. Nucleic acids research *43*, D6-17.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics : MCP *13*, 2513-2526.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

Cox, J., and Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinformatics *13 Suppl 16*, S12.

Crowley, P.B., and Golovin, A. (2005). Cation-pi interactions in protein-protein interfaces. Proteins *59*, 231-239.

DeRenzo, C., Reese, K.J., and Seydoux, G. (2003). Exclusion of germ plasm proteins from somatic lineages by cullin-dependent degradation. Nature *424*, 685-689.

Dinkel, H., Van Roey, K., Michael, S., Davey, N.E., Weatheritt, R.J., Born, D., Speck, T., Kruger, D., Grebnev, G., Kuban, M.*, et al.* (2014). The eukaryotic linear motif resource ELM: 10 years and counting. Nucleic acids research *42*, D259-266.

Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433-3434.

Dosztanyi, Z., Meszaros, B., and Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics *25*, 2745-2746.

Eickbush, T.H., and Moudrianakis, E.N. (1978). The histone core complex: an octamer assembled by two sets of protein-protein interactions. Biochemistry *17*, 4955-4964.

Ewing, R.M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M.*, et al.* (2007). Large-scale mapping of human protein-protein interactions by mass spectrometry. Molecular systems biology *3*, 89.

Falcon, S., and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. Bioinformatics *23*, 257-258.

Fernandez, A.G., Gunsalus, K.C., Huang, J., Chuang, L.S., Ying, N., Liang, H.L., Tang, C., Schetter, A.J., Zegar, C., Rual, J.F.*, et al.* (2005). New genes with roles in the C. elegans embryo revealed using RNAi of ovary-enriched ORFeome clones. Genome Res *15*, 250-259.

Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325-330.

Fraser, H.B., and Plotkin, J.B. (2007). Using protein complexes to predict phenotypic effects of gene mutation. Genome biology *8*, R252.

Fridy, P.C., Li, Y., Keegan, S., Thompson, M.K., Nudelman, I., Scheid, J.F., Oeffinger, M., Nussenzweig, M.C., Fenyo, D., Chait, B.T.*, et al.* (2014). A robust pipeline for rapid production of versatile nanobody repertoires. Nature methods *11*, 1253-1260.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene *313*, 17-42.

Gallie, D.R. (2002). Protein-protein interactions required during translation. Plant molecular biology *50*, 949-970.

Gallo, C.M., Wang, J.T., Motegi, F., and Seydoux, G. (2010). Cytoplasmic partitioning of P granule components is not required to specify the germline in C. elegans. Science *330*, 1685-1689.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M.*, et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Gingras, A.C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nature reviews Molecular cell biology *8*, 645-654.

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E.*, et al.* (2003). A protein interaction map of Drosophila melanogaster. Science *302*, 1727-1736.

Goldstein, B., and Hird, S.N. (1996). Specification of the anteroposterior axis in Caenorhabditis elegans. Development *122*, 1467-1474.

Gonczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S.J., Copley, R.R., Duperon, J., Oegema, J., Brehm, M., Cassin, E.*, et al.* (2000). Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. Nature *408*, 331-336.

Grossmann, J., Roschitzki, B., Panse, C., Fortes, C., Barkow-Oesterreicher, S., Rutishauser, D., and Schlapbach, R. (2010). Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. J Proteomics *73*, 1740-1746.

Gu, Z.C., and Enenkel, C. (2014). Proteasome assembly. Cellular and molecular life sciences : CMLS *71*, 4729-4745.

Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G*., et al.* (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell *149*, 768-779.

Hanack, C., Moroni, M., Lima, W.C., Wende, H., Kirchner, M., Adelfinger, L., Schrenk-Siemens, K., Tappe-Theodor, A., Wetzel, C., Kuich, P.H*., et al.* (2015). GABA Blocks Pathological but Not Acute TRPV1 Pain Signals. Cell *160*, 759-770.

Hanazawa, M., Yonetani, M., and Sugimoto, A. (2011). PGL proteins self associate and bind RNPs to mediate germ granule assembly in C. elegans. The Journal of cell biology *192*, 929-937.

Hird, S.N., Paulsen, J.E., and Strome, S. (1996). Segregation of germ granules in living Caenorhabditis elegans embryos: cell-type-specific mechanisms for cytoplasmic localisation. Development *122*, 1303-1312.

Hoege, C., Constantinescu, A.T., Schwager, A., Goehring, N.W., Kumar, P., and Hyman, A.A. (2010). LGL can partition the cortex of one-cell Caenorhabditis elegans embryos into two domains. Current biology : CB *20*, 1296-1303.

Hoege, C., and Hyman, A.A. (2013). Principles of PAR polarity in Caenorhabditis elegans embryos. Nature reviews Molecular cell biology *14*, 315-322.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.

Huang, N.N., Mootz, D.E., Walhout, A.J., Vidal, M., and Hunter, C.P. (2002). MEX-3 interacting proteins link cell polarity to asymmetric gene expression in Caenorhabditis elegans. Development *129*, 747-759.

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. The Journal of cell biology *189*, 739-754.

Hubner, N.C., and Mann, M. (2011). Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods *53*, 453-459.

Hyman, A.A., Weber, C.A., and Julicher, F. (2014). Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol *30*, 39-58.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Molecular & cellular proteomics : MCP *4*, 1265-1272.

Ivanov, A.A., Khuri, F.R., and Fu, H. (2013). Targeting protein-protein interactions as an anticancer strategy. Trends in pharmacological sciences *34*, 393-400.

Jones, N., Blasutig, I.M., Eremina, V., Ruston, J.M., Bladt, F., Li, H., Huang, H., Larose, L., Li, S.S., Takano, T*., et al.* (2006). Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. Nature *440*, 818-823.

Jungkamp, A.C., Stoeckius, M., Mecenas, D., Grun, D., Mastrobuoni, G., Kempa, S., and Rajewsky, N. (2011). In vivo and transcriptome-wide identification of RNA binding protein target sites. Molecular cell *44*, 828-840.

Kaitna, S., Mendoza, M., Jantsch-Plunger, V., and Glotzer, M. (2000). Incenp and an aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. Current biology : CB *10*, 1172-1181.

Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G., and Ahringer, J. (2001). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. Genome biology *2*, RESEARCH0002.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J.*, et al.* (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753-767.

Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. Cell *94*, 635-645.

Kedersha, N., Ivanov, P., and Anderson, P. (2013). Stress granules and cell signaling: more than just a passing phase? Trends in biochemical sciences *38*, 494-506.

Keilhauer, E.C., Hein, M.Y., and Mann, M. (2015). Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). Molecular & cellular proteomics : MCP *14*, 120-135.

Kemphues, K.J., Priess, J.R., Morton, D.G., and Cheng, N.S. (1988). Identification of genes required for cytoplasmic localization in early C. elegans embryos. Cell *52*, 311-320.

Kiebler, M.A., and Bassell, G.J. (2006). Neuronal RNA granules: movers and makers. Neuron *51*, 685-690.

Krieg, C., Cole, T., Deppe, U., Schierenberg, E., Schmitt, D., Yoder, B., and con Ehrenstein, G. (1978). The cellular anatomy of embryos of the nematode Caenorhabditis elegans. Analysis and reconstruction of serial section electron micrographs. Developmental biology *65*, 193-215.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049-1060.

LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C.*, et al.* (2005). A protein interaction network of the malaria parasite Plasmodium falciparum. Nature *438*, 103-107.

Leacock, S.W., and Reinke, V. (2008). MEG-1 and MEG-2 are embryo-specific P-granule components required for germline development in Caenorhabditis elegans. Genetics *178*, 295-306.

Leckband, D. (2000). Measuring the forces that control protein interactions. Annu Rev Biophys Biomol Struct *29*, 1-26.

Li, P., Banjade, S., Cheng, H.C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F.*, et al.* (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature *483*, 336-340.

Li, S., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.O., Han, J.D., Chesneau, A., Hao, T.*, et al.* (2004). A map of the interactome network of the metazoan C. elegans. Science *303*, 540-543.

Li, W. (2012). Volcano plots in analyzing differential expressions with mRNA microarrays. Journal of bioinformatics and computational biology *10*, 1231003.

Li, X., Foley, E.A., Kawashima, S.A., Molloy, K.R., Li, Y., Chait, B.T., and Kapoor, T.M. (2013a). Examining post-translational modification-mediated protein-protein interactions using a chemical proteomics approach. Protein Sci *22*, 287-295.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013b). Stress granules as crucibles of ALS pathogenesis. The Journal of cell biology *201*, 361-372.

Lowe, N., Rees, J.S., Roote, J., Ryder, E., Armean, I.M., Johnson, G., Drummond, E., Spriggs, H., Drummond, J., Magbanua, J.P*., et al.* (2014). Analysis of the expression patterns, subcellular localisations and interaction partners of Drosophila proteins using a pigP protein trap library. Development *141*, 3994-4005.

Lundby, A., Rossin, E.J., Steffensen, A.B., Acha, M.R., Newton-Cheh, C., Pfeufer, A., Lynch, S.N., Consortium, Q.T.I.I.G., Olesen, S.P., Brunak, S*., et al.* (2014). Annotation of loci from genome-wide association studies using tissue-specific quantitative interaction proteomics. Nature methods *11*, 868-874.

Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W., Ding, C., Shi, Y., Yucer, N., Krenciute, G*., et al.* (2011). Analysis of the human endogenous coregulator complexome. Cell *145*, 787-799.

Marcello, M.R., and Singson, A. (2010). Fertilization and the oocyte-to-embryo transition in C. elegans. BMB Rep *43*, 389-399.

McCraith, S., Holtzman, T., Moss, B., and Fields, S. (2000). Genome-wide analysis of vaccinia virus protein-protein interactions. Proceedings of the National Academy of Sciences of the United States of America *97*, 4879-4884.

Mello, C.C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J.R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. Nature *382*, 710-712.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Current biology : CB *18*, 1476-1482.

Meszaros, B., Simon, I., and Dosztanyi, Z. (2009). Prediction of protein binding regions in disordered proteins. PLoS Comput Biol *5*, e1000376.

Mishima, M., Kaitna, S., and Glotzer, M. (2002). Central spindle assembly and cytokinesis require a kinesin-like protein/RhoGAP complex with microtubule bundling activity. Developmental cell *2*, 41-54.

Mousson, F., Kolkman, A., Pijnappel, W.W., Timmers, H.T., and Heck, A.J. (2008). Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. Molecular & cellular proteomics : MCP *7*, 845-852.

Nishi, Y., and Lin, R. (2005). DYRK2 and GSK-3 phosphorylate and promote the timely degradation of OMA-1, a key regulator of the oocyte-to-embryo transition in C. elegans. Developmental biology *288*, 139-149.

Nishi, Y., Rogers, E., Robertson, S.M., and Lin, R. (2008). Polo kinases regulate C. elegans embryonic polarity via binding to DYRK2-primed MEX-5 and MEX-6. Development *135*, 687-697.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D*., et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Molecular cell *57*, 936-947.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & cellular proteomics : MCP *1*, 376-386.

Ozbabacan, S.E., Engin, H.B., Gursoy, A., and Keskin, O. (2011). Transient protein-protein interactions. Protein engineering, design & selection : PEDS *24*, 635-648.

Pagano, J.M., Farley, B.M., McCoig, L.M., and Ryder, S.P. (2007). Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. The Journal of biological chemistry *282*, 8883-8894.

Pang, K.M., Ishidate, T., Nakamura, K., Shirayama, M., Trzepacz, C., Schubert, C.M., Priess, J.R., and Mello, C.C. (2004). The minibrain kinase homolog, mbk-2, is required for spindle positioning and asymmetric cell division in early C. elegans embryos. Developmental biology *265*, 127-139.

Parrish, J.R., Gulyas, K.D., and Finley, R.L., Jr. (2006). Yeast two-hybrid contributions to interactome mapping. Current opinion in biotechnology *17*, 387-393.

Paul, F.E., Hosp, F., and Selbach, M. (2011). Analyzing protein-protein interactions by quantitative mass spectrometry. Methods *54*, 387-395.

Pellettieri, J., Reinke, V., Kim, S.K., and Seydoux, G. (2003). Coordinate activation of maternal protein degradation during the egg-to-embryo transition in C. elegans. Developmental cell *5*, 451-462.

Pinna, L.A., and Ruzzene, M. (1996). How do protein kinases recognize their substrates? Biochim Biophys Acta *1314*, 191-225.

Pitt, J.N., Schisa, J.A., and Priess, J.R. (2000). P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. Developmental biology *219*, 315-333.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218-229.

Raich, W.B., Moorman, C., Lacefield, C.O., Lehrer, J., Bartsch, D., Plasterk, R.H., Kandel, E.R., and Hobert, O. (2003). Characterization of Caenorhabditis elegans homologs of the Down syndrome candidate gene DYRK1A. Genetics *163*, 571-580.

Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V*., et al.* (2001). The protein-protein interaction map of Helicobacter pylori. Nature *409*, 211-215.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc *2*, 1896-1906.

Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res *12*, 1231-1245.

Rees, J.S., Lowe, N., Armean, I.M., Roote, J., Johnson, G., Drummond, E., Spriggs, H., Ryder, E., Russell, S., St Johnston, D*., et al.* (2011). In vivo analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. Molecular & cellular proteomics : MCP *10*, M110 002386.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet *16*, 276-277.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature biotechnology *17*, 1030-1032.

Rinner, O., Mueller, L.N., Hubalek, M., Muller, M., Gstaiger, M., and Aebersold, R. (2007). An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. Nature biotechnology *25*, 345-352.

Robertson, S., and Lin, R. (2013). The oocyte-to-embryo transition. Advances in experimental medicine and biology *757*, 351-372.

Ross, C.A., and Poirier, M.A. (2005). Opinion: What is the role of protein aggregation in neurodegeneration? Nature reviews Molecular cell biology *6*, 891-898.

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C., and Leonhardt, H. (2008). A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Molecular & cellular proteomics : MCP *7*, 282-289.

Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N*., et al.* (2005). Towards a proteome-scale map of the human protein-protein interaction network. Nature *437*, 1173-1178.

Sarov, M., Murray, J.I., Schanze, K., Pozniakovski, A., Niu, W., Angermann, K., Hasse, S., Rupprecht, M., Vinis, E., Tinney, M*., et al.* (2012). A Genome-Scale Resource for In Vivo Tag-Based Protein Function Exploration in C. elegans. Cell *150*, 855-866.

Scheckel, C., Gaidatzis, D., Wright, J.E., and Ciosk, R. (2012). Genome-wide analysis of GLD-1-mediated mRNA regulation suggests a role in mRNA storage. PLoS genetics *8*, e1002742.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature methods *9*, 671-675.

Schonegg, S., Constantinescu, A.T., Hoege, C., and Hyman, A.A. (2007). The Rho GTPase-activating proteins RGA-3 and RGA-4 are required to set the initial size of PAR domains in Caenorhabditis elegans one-cell embryos. Proceedings of the National Academy of Sciences of the United States of America *104*, 14976-14981.

Schubert, C.M., Lin, R., de Vries, C.J., Plasterk, R.H., and Priess, J.R. (2000). MEX-5 and MEX-6 function to establish soma/germline asymmetry in early C. elegans embryos. Molecular cell *5*, 671-682.

Selbach, M., and Mann, M. (2006). Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nature methods *3*, 981-983.

Sengupta, M.S., Low, W.Y., Patterson, J.R., Kim, H.M., Traven, A., Beilharz, T.H., Colaiacovo, M.P., Schisa, J.A., and Boag, P.R. (2013). ifet-1 is a broad-scale translational repressor required for normal P granule formation in C. elegans. Journal of cell science *126*, 850-859.

Seydoux, G., and Fire, A. (1994). Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. Development *120*, 2823-2834.

Simonis, N., Rual, J.F., Carvunis, A.R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Sahalie, J.M., Venkatesan, K., Gebreab, F.*, et al.* (2009). Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome network. Nature methods *6*, 47-54.

Simpson-Holley, M., Kedersha, N., Dower, K., Rubins, K.H., Anderson, P., Hensley, L.E., and Connor, J.H. (2011). Formation of antiviral cytoplasmic granules during orthopoxvirus infection. J Virol *85*, 1581-1593.

Smits, A.H., Jansen, P.W., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic acids research *41*, e28.

Somasekharan, S.P., El-Naggar, A., Leprivier, G., Cheng, H., Hajee, S., Grunewald, T.G., Zhang, F., Ng, T., Delattre, O., Evdokimova, V.*, et al.* (2015). YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. The Journal of cell biology *208*, 913-929.

Sonnichsen, B., Koski, L.B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A.M., Artelt, J., Bettencourt, P., Cassin, E.*, et al.* (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature *434*, 462-469.

Speese, S.D., Ashley, J., Jokhi, V., Nunnari, J., Barria, R., Li, Y., Ataman, B., Koon, A., Chang, Y.T., Li, Q.*, et al.* (2012). Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. Cell *149*, 832-846.

Squirrell, J.M., Eggers, Z.T., Luedke, N., Saari, B., Grimson, A., Lyons, G.E., Anderson, P., and White, J.G. (2006). CAR-1, a protein that localizes with the mRNA decapping component DCAP-1, is required for cytokinesis and ER organization in Caenorhabditis elegans embryos. Molecular biology of the cell *17*, 336-344.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S.*, et al.* (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell *122*, 957-968.

Stitzel, M.L., Pellettieri, J., and Seydoux, G. (2006). The C. elegans DYRK Kinase MBK-2 Marks Oocyte Proteins for Degradation in Response to Meiotic Maturation. Current biology : CB *16*, 56-62.

Stoeckius, M., Grun, D., Kirchner, M., Ayoub, S., Torti, F., Piano, F., Herzog, M., Selbach, M., and Rajewsky, N. (2014). Global characterization of the oocyte-to-embryo transition in Caenorhabditis elegans uncovers a novel mRNA clearance mechanism. The EMBO journal *33*, 1751-1766.

Stroedicke, M., Bounab, Y., Strempel, N., Klockmeier, K., Yigit, S., Friedrich, R.P., Chaurasia, G., Li, S., Hesse, F., Riechers, S.P.*, et al.* (2015). Systematic interaction network filtering identifies CRMP1 as a novel suppressor of huntingtin misfolding and neurotoxicity. Genome Res *25*, 701-713.

Strome, S., and Wood, W.B. (1982). Immunofluorescence visualization of germ-line-specific cytoplasmic granules in embryos, larvae, and adults of Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *79*, 1558-1562.

Strome, S., and Wood, W.B. (1983). Generation of asymmetry and segregation of germ-line granules in early C. elegans embryos. Cell *35*, 15-25.

Subbotin, R.I., and Chait, B.T. (2014). A pipeline for determining protein-protein interactions and proximities in the cellular milieu. Molecular & cellular proteomics : MCP *13*, 2824-2835.

Suchanek, M., Radzikowska, A., and Thiele, C. (2005). Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nature methods *2*, 261-267.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Developmental biology *100*, 64-119.

Suzuki, A., and Ohno, S. (2006). The PAR-aPKC system: lessons in polarity. Journal of cell science *119*, 979-987.

Tabara, H., Hill, R.J., Mello, C.C., Priess, J.R., and Kohara, Y. (1999). pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans. Development *126*, 1-11.

Takeda, K., Watanabe, C., Qadota, H., Hanazawa, M., and Sugimoto, A. (2008). Efficient production of monoclonal antibodies recognizing specific structures in Caenorhabditis elegans embryos using an antigen subtraction method. Genes to cells : devoted to molecular & cellular mechanisms *13*, 653-665.

Tian, Y., Li, Z., Hu, W., Ren, H., Tian, E., Zhao, Y., Lu, Q., Huang, X., Yang, P., Li, X*., et al.* (2010). C. elegans screen identifies autophagy genes specific to multicellular organisms. Cell *141*, 1042-1055.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America *98*, 5116-5121.

Tzur, Y.B., Friedland, A.E., Nadarajan, S., Church, G.M., Calarco, J.A., and Colaiacovo, M.P. (2013). Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. Genetics *195*, 1181-1185.

Uetz, P., Dong, Y.A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S.V., Roupelieva, M., Rose, D., Fossum, E*., et al.* (2006). Herpesviral protein networks and their interaction with the human proteome. Science *311*, 239-242.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P*., et al.* (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature *403*, 623-627.

Updike, D., and Strome, S. (2010). P granule assembly and function in Caenorhabditis elegans germ cells. Journal of andrology *31*, 53-60.

Updike, D.L., Knutson, A.K., Egelhofer, T.A., Campbell, A.C., and Strome, S. (2014). Germ-granule components prevent somatic development in the C. elegans germline. Current biology : CB *24*, 970-975.

Updike, D.L., and Strome, S. (2009). A genomewide RNAi screen for genes that affect the stability, distribution and function of P granules in Caenorhabditis elegans. Genetics *183*, 1397-1419.

Vermeulen, M., Hubner, N.C., and Mann, M. (2008). High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. Current opinion in biotechnology *19*, 331-337.

Voronina, E., and Seydoux, G. (2010). The C. elegans homolog of nucleoporin Nup98 is required for the integrity and function of germline P granules. Development *137*, 1441-1450.

Voronina, E., Seydoux, G., Sassone-Corsi, P., and Nagamori, I. (2011). RNA granules in germ cells. Cold Spring Harbor perspectives in biology *3*.

Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N., and Vidal, M. (2000). Protein interaction mapping in C. elegans using proteins involved in vulval development. Science *287*, 116-122.

Wang, J.T., Smith, J., Chen, B., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically-disordered proteins in. eLife *3*.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and protein in phase. Cell *149*, 1188-1191.

Whitty, A. (2008). Cooperativity and biological complexity. Nat Chem Biol *4*, 435-439.

Wippich, F., Bodenmiller, B., Trajkovska, M.G., Wanka, S., Aebersold, R., and Pelkmans, L. (2013). Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. Cell *152*, 791-805.

Wolf, N., Priess, J., and Hirsh, D. (1983). Segregation of germline granules in early embryos of Caenorhabditis elegans: an electron microscopic analysis. J Embryol Exp Morphol *73*, 297-306.

Wootton, J.C. (1994). Non-globular domains in protein sequences: automated segmentation using complexity measures. Comput Chem *18*, 269-285.

Wootton, J.C., and Federhen, S. (1996). Analysis of compositionally biased regions in sequence databases. Methods in enzymology *266*, 554-571.

Wylie, C. (2000). Germ cells. Curr Opin Genet Dev *10*, 410-413.

Yamamoto, I., Kosinski, M.E., and Greenstein, D. (2006). Start me up: cell signaling and the journey from oocyte to embryo in C. elegans. Dev Dyn *235*, 571-585.

Yeung, Y.G., and Stanley, E.R. (2009). A solution for stripping antibodies from polyvinylidene fluoride immunoblots for multiple reprobing. Anal Biochem *389*, 89-91.

Yuet, K.P., Doma, M.K., Ngo, J.T., Sweredoski, M.J., Graham, R.L., Moradian, A., Hess, S., Schuman, E.M., Sternberg, P.W., and Tirrell, D.A. (2015). Cell-specific proteomic analysis in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *112*, 2705-2710.

Zeiser, E., Frokjaer-Jensen, C., Jorgensen, E., and Ahringer, J. (2011). MosSCI and gateway compatible plasmid toolkit for constitutive and inducible expression of transgenes in the C. elegans germline. PloS one *6*, e20082.

Zhang, Y., Yan, L., Zhou, Z., Yang, P., Tian, E., Zhang, K., Zhao, Y., Li, Z., Song, B., Han, J.*, et al.* (2009). SEPA-1 mediates the specific recognition and degradation of P granule components by autophagy in C. elegans. Cell *136*, 308-321.

Zubarev, R.A. (2013). The challenge of the proteome dynamic range and its implications for in-depth proteomics. Proteomics *13*, 723-726.

VII Supplementary Information

VII.1. Abbreviations

Abbreviation	Full name
A	Ampere
ABC	ammonium bicarbonate
cDNA	complementary deoxyribonucleic acid
CID	collision-induced dissociation
DMEM	Dulbecco's Modified Eagle Medium
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DYRK	dual-specificity tyrosine phosphorylation-regulated kinase
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EGTA	ethylene glycol tetraacetic acid
emPAI	exponentially modified protein abundance index
FDR	false discovery rate
GFP	green fluorescent protein
GO	gene ontology
HCD	higher energy collision dissociation
HEK	human embryonic kidney
hr	hour
IP	immunoprecipitation
IPTG	isopropyl β-D-1-thiogalactopyranoside
IVI	*in vivo* interactome
LB	lysogeny broth
LCI	literature-curated interaction
LC-MS	liquid chromatography-mass spectrometry
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LFQ	label-free quantification
LTQ	linear trap quadrupole
Lys-C	lysyl endopeptidase
min	minute
MOPS	3-morpholinopropane-1-sulfonic acid
MS	mass spectrometry
ms	millisecond
NGM	nematode growth media

Abbreviation	Full name
PAI	protein abundance index
PBS	phosphate-buffered saline
PBST	phosphate-buffered saline Tween-20
PCR	polymerase chain reaction
PEI	polyethylenimine
PP2A	protein phosphatase 2A
PRM	proline-rich motif
PTM	post-translational modification
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
s	second
SAM	statistical analysis of microarray
SDS	sodium dodecyl sulfate
SH3	SRC Homology 3
SILAC	stable isotope labeling by amino acids in cell culture
SOC	super optimal broth with catabolite repression
StageTip	stop and go extraction tip
TAE	Tris-acetate-EDTA
TAP	tandem affinity purification
TBST	Tris-buffered saline Tween-20
TPR	true positive rate
V	volt
WI8	Worm Interactome version 8
XIC	extracted ion chromatogram

VII.2. Supplementary figure

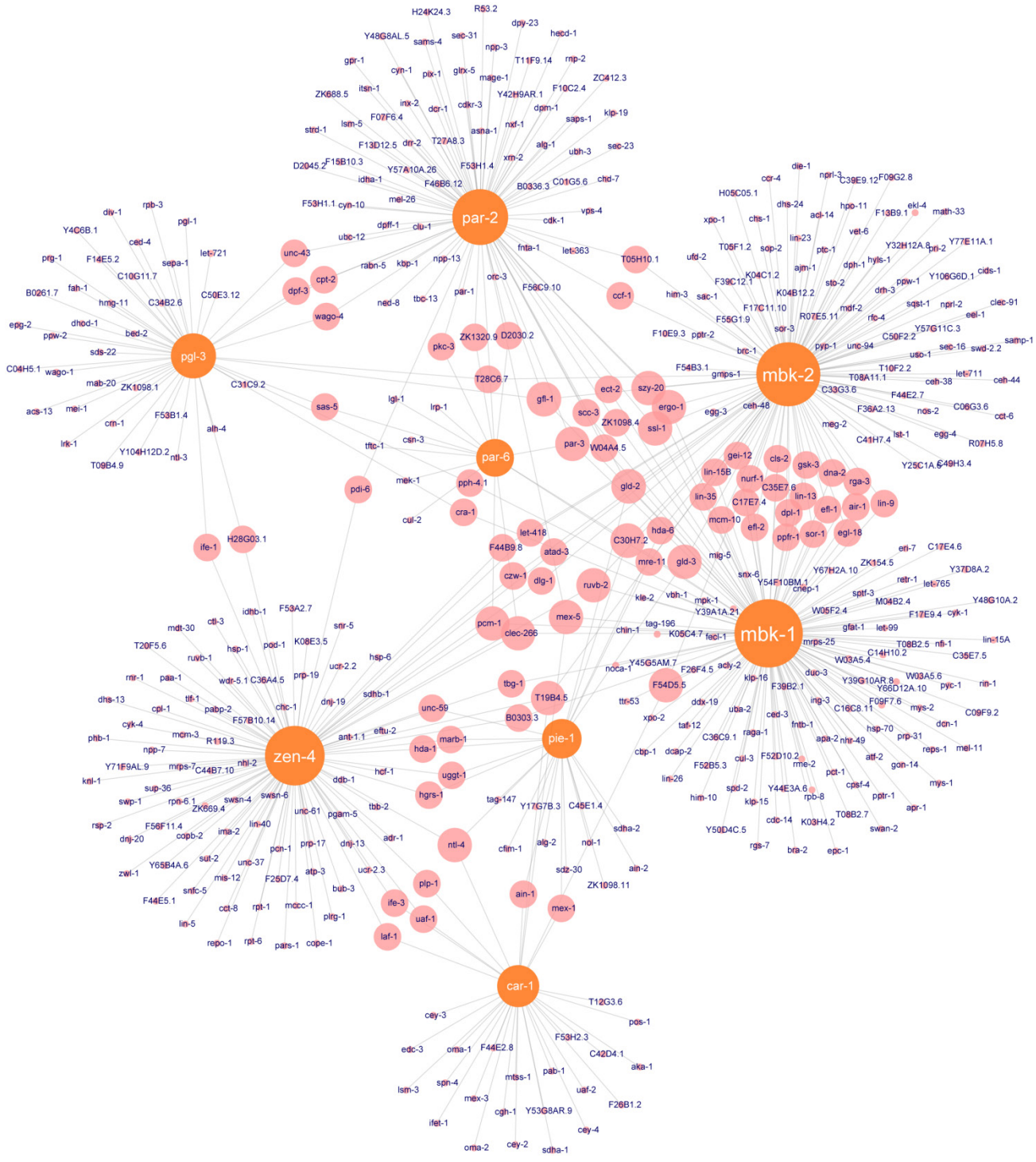

Fig. VII.1 | Embryo *in vivo* interactome (IVI) network annotated with protein names. Bait (orange) and prey (pink) proteins are shown as nodes sized proportionally to their degree distribution.

VII.3. Supplementary table

Table VII.1 | Complete list of embryo *in vivo* interactome data

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
Y18D10A.17	WBGene00012484	car-1	D1022.7	WBGene00000101	aka-1
Y18D10A.17	WBGene00012484	car-1	F46F11.2	WBGene00000473	cey-2
Y18D10A.17	WBGene00012484	car-1	M01E11.5	WBGene00000474	cey-3
Y18D10A.17	WBGene00012484	car-1	Y39A1C.3	WBGene00000475	cey-4
Y18D10A.17	WBGene00012484	car-1	C07H6.5	WBGene00000479	cgh-1
Y18D10A.17	WBGene00012484	car-1	B0348.6	WBGene00002061	ife-3
Y18D10A.17	WBGene00012484	car-1	Y71H2AM.19	WBGene00002244	laf-1
Y18D10A.17	WBGene00012484	car-1	Y62E10A.12	WBGene00003077	lsm-3
Y18D10A.17	WBGene00012484	car-1	W03C9.7	WBGene00003228	mex-1
Y18D10A.17	WBGene00012484	car-1	F53G12.5	WBGene00003229	mex-3
Y18D10A.17	WBGene00012484	car-1	W02A2.7	WBGene00003230	mex-5
Y18D10A.17	WBGene00012484	car-1	C49H3.5	WBGene00003827	ntl-4
Y18D10A.17	WBGene00012484	car-1	C09G9.6	WBGene00003864	oma-1
Y18D10A.17	WBGene00012484	car-1	ZC513.6	WBGene00003865	oma-2
Y18D10A.17	WBGene00012484	car-1	Y106G6H.2	WBGene00003902	pab-1
Y18D10A.17	WBGene00012484	car-1	F45E4.2	WBGene00004046	plp-1
Y18D10A.17	WBGene00012484	car-1	F52E1.1	WBGene00004078	pos-1
Y18D10A.17	WBGene00012484	car-1	F56F3.1	WBGene00004132	ifet-1
Y18D10A.17	WBGene00012484	car-1	ZC404.8	WBGene00004984	spn-4
Y18D10A.17	WBGene00012484	car-1	Y92C3B.2	WBGene00006697	uaf-1
Y18D10A.17	WBGene00012484	car-1	Y116A8C.35	WBGene00006698	uaf-2
Y18D10A.17	WBGene00012484	car-1	F53H2.3	WBGene00010006	F53H2.3
Y18D10A.17	WBGene00012484	car-1	F54D5.5	WBGene00010051	F54D5.5
Y18D10A.17	WBGene00012484	car-1	R05D11.8	WBGene00011036	edc-3
Y18D10A.17	WBGene00012484	car-1	T12G3.6	WBGene00011741	T12G3.6
Y18D10A.17	WBGene00012484	car-1	C03G5.1	WBGene00015391	sdha-1
Y18D10A.17	WBGene00012484	car-1	C06G1.4	WBGene00015547	ain-1
Y18D10A.17	WBGene00012484	car-1	C42D4.1	WBGene00016594	C42D4.1
Y18D10A.17	WBGene00012484	car-1	F26B1.2	WBGene00017816	F26B1.2
Y18D10A.17	WBGene00012484	car-1	F44E2.8	WBGene00018421	F44E2.8
Y18D10A.17	WBGene00012484	car-1	PAR2.1	WBGene00019800	mtss-1
Y18D10A.17	WBGene00012484	car-1	T19B4.5	WBGene00020558	T19B4.5
Y18D10A.17	WBGene00012484	car-1	Y53G8AR.9	WBGene00021816	Y53G8AR.9
T04C10.1	WBGene00003149	mbk-1	K07C11.2	WBGene00000098	air-1
T04C10.1	WBGene00003149	mbk-1	K04G2.8	WBGene00000156	apr-1
T04C10.1	WBGene00003149	mbk-1	T20B5.1	WBGene00000161	apa-2
T04C10.1	WBGene00003149	mbk-1	K08F8.2	WBGene00000220	atf-2
T04C10.1	WBGene00003149	mbk-1	R10E11.1	WBGene00000366	cbp-1
T04C10.1	WBGene00003149	mbk-1	C17G10.4	WBGene00000383	cdc-14
T04C10.1	WBGene00003149	mbk-1	C48D1.2	WBGene00000417	ced-3
T04C10.1	WBGene00003149	mbk-1	R107.6	WBGene00000549	cls-2
T04C10.1	WBGene00003149	mbk-1	Y108G3AL.1	WBGene00000838	cul-3
T04C10.1	WBGene00003149	mbk-1	F11H8.4	WBGene00000872	cyk-1
T04C10.1	WBGene00003149	mbk-1	F43G6.1	WBGene00001016	dna-2
T04C10.1	WBGene00003149	mbk-1	T23G7.1	WBGene00001061	dpl-1
T04C10.1	WBGene00003149	mbk-1	Y106G6H.12	WBGene00001112	duo-3
T04C10.1	WBGene00003149	mbk-1	Y102A5C.18	WBGene00001161	efl-1
T04C10.1	WBGene00003149	mbk-1	Y48C3A.17	WBGene00001162	efl-2
T04C10.1	WBGene00003149	mbk-1	F55A8.1	WBGene00001186	egl-18

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
T04C10.1	WBGene00003149	mbk-1	F52D2.4	WBGene00001569	gei-12
T04C10.1	WBGene00003149	mbk-1	M04B2.3	WBGene00001585	gfl-1
T04C10.1	WBGene00003149	mbk-1	ZC308.1	WBGene00001596	gld-2
T04C10.1	WBGene00003149	mbk-1	T07F8.3	WBGene00001597	gld-3
T04C10.1	WBGene00003149	mbk-1	Y18D10A.5	WBGene00001746	gsk-3
T04C10.1	WBGene00003149	mbk-1	R12B2.4	WBGene00001869	him-10
T04C10.1	WBGene00003149	mbk-1	C12C8.1	WBGene00002026	hsp-70
T04C10.1	WBGene00003149	mbk-1	Y48G1A.5	WBGene00002079	xpo-2
T04C10.1	WBGene00003149	mbk-1	F44C4.4	WBGene00002148	gon-14
T04C10.1	WBGene00003149	mbk-1	M01E11.6	WBGene00002225	klp-15
T04C10.1	WBGene00003149	mbk-1	C41G7.2	WBGene00002226	klp-16
T04C10.1	WBGene00003149	mbk-1	T19E10.1	WBGene00002297	ect-2
T04C10.1	WBGene00003149	mbk-1	K08E7.3	WBGene00002368	let-99
T04C10.1	WBGene00003149	mbk-1	F20H11.2	WBGene00002889	let-765
T04C10.1	WBGene00003149	mbk-1	ZK637.7	WBGene00002998	lin-9
T04C10.1	WBGene00003149	mbk-1	C03B8.4	WBGene00003002	lin-13
T04C10.1	WBGene00003149	mbk-1	F18A1.2	WBGene00003012	lin-26
T04C10.1	WBGene00003149	mbk-1	C32F10.2	WBGene00003020	lin-35
T04C10.1	WBGene00003149	mbk-1	C06C3.1	WBGene00003196	mel-11
T04C10.1	WBGene00003149	mbk-1	W02A2.7	WBGene00003230	mex-5
T04C10.1	WBGene00003149	mbk-1	T05C12.6	WBGene00003241	mig-5
T04C10.1	WBGene00003149	mbk-1	F43C1.2	WBGene00003401	mpk-1
T04C10.1	WBGene00003149	mbk-1	F52G2.1	WBGene00003582	dcap-2
T04C10.1	WBGene00003149	mbk-1	ZK1290.4	WBGene00003592	nfi-1
T04C10.1	WBGene00003149	mbk-1	K10C3.6	WBGene00003639	nhr-49
T04C10.1	WBGene00003149	mbk-1	F54E7.3	WBGene00003918	par-3
T04C10.1	WBGene00003149	mbk-1	T26E3.3	WBGene00003921	par-6
T04C10.1	WBGene00003149	mbk-1	C10F3.5	WBGene00003954	pcm-1
T04C10.1	WBGene00003149	mbk-1	C07G1.3	WBGene00003961	pct-1
T04C10.1	WBGene00003149	mbk-1	Y75B8A.30	WBGene00004085	pph-4.1
T04C10.1	WBGene00003149	mbk-1	C18E9.3	WBGene00004105	szy-20
T04C10.1	WBGene00003149	mbk-1	D2023.2	WBGene00004258	pyc-1
T04C10.1	WBGene00003149	mbk-1	F56B6.2	WBGene00004350	rgs-7
T04C10.1	WBGene00003149	mbk-1	T11F8.3	WBGene00004374	rme-2
T04C10.1	WBGene00003149	mbk-1	F18E2.3	WBGene00004738	scc-3
T04C10.1	WBGene00003149	mbk-1	F32H2.3	WBGene00004953	spd-2
T04C10.1	WBGene00003149	mbk-1	Y56A3A.4	WBGene00006396	taf-12
T04C10.1	WBGene00003149	mbk-1	T24F1.1	WBGene00006414	raga-1
T04C10.1	WBGene00003149	mbk-1	F23B12.6	WBGene00006465	fntb-1
T04C10.1	WBGene00003149	mbk-1	F58A4.8	WBGene00006540	tbg-1
T04C10.1	WBGene00003149	mbk-1	W02A11.4	WBGene00006700	uba-2
T04C10.1	WBGene00003149	mbk-1	Y54E10A.9	WBGene00006888	vbh-1
T04C10.1	WBGene00003149	mbk-1	Y111B2A.22	WBGene00007027	ssl-1
T04C10.1	WBGene00003149	mbk-1	VC5.4	WBGene00007029	mys-1
T04C10.1	WBGene00003149	mbk-1	Y111B2A.11	WBGene00007030	epc-1
T04C10.1	WBGene00003149	mbk-1	F41E6.6	WBGene00007055	tag-196
T04C10.1	WBGene00003149	mbk-1	B0365.1	WBGene00007150	acly-2
T04C10.1	WBGene00003149	mbk-1	C09F9.2	WBGene00007479	C09F9.2
T04C10.1	WBGene00003149	mbk-1	C14H10.2	WBGene00007592	C14H10.2
T04C10.1	WBGene00003149	mbk-1	C17E4.6	WBGene00007645	C17E4.6
T04C10.1	WBGene00003149	mbk-1	C48G7.3	WBGene00008183	rin-1
T04C10.1	WBGene00003149	mbk-1	F07A11.2	WBGene00008546	gfat-1
T04C10.1	WBGene00003149	mbk-1	F16A11.3	WBGene00008878	ppfr-1

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
T04C10.1	WBGene00003149	mbk-1	F26H11.2	WBGene00009180	nurf-1
T04C10.1	WBGene00003149	mbk-1	F39B2.1	WBGene00009553	F39B2.1
T04C10.1	WBGene00003149	mbk-1	F52B5.3	WBGene00009922	F52B5.3
T04C10.1	WBGene00003149	mbk-1	F52D10.2	WBGene00009930	F52D10.2
T04C10.1	WBGene00003149	mbk-1	F53C11.7	WBGene00009976	swan-2
T04C10.1	WBGene00003149	mbk-1	F54D5.5	WBGene00010051	F54D5.5
T04C10.1	WBGene00003149	mbk-1	H38K22.2	WBGene00010428	dcn-1
T04C10.1	WBGene00003149	mbk-1	K03D10.3	WBGene00010537	mys-2
T04C10.1	WBGene00003149	mbk-1	K03H4.2	WBGene00010549	K03H4.2
T04C10.1	WBGene00003149	mbk-1	K05C4.7	WBGene00010583	K05C4.7
T04C10.1	WBGene00003149	mbk-1	K07G5.6	WBGene00010645	fecl-1
T04C10.1	WBGene00003149	mbk-1	M04B2.4	WBGene00010847	M04B2.4
T04C10.1	WBGene00003149	mbk-1	T07D4.4	WBGene00011580	ddx-19
T04C10.1	WBGene00003149	mbk-1	T09E8.1	WBGene00011647	noca-1
T04C10.1	WBGene00003149	mbk-1	T14G10.3	WBGene00011773	ttr-53
T04C10.1	WBGene00003149	mbk-1	W04A4.5	WBGene00012234	W04A4.5
T04C10.1	WBGene00003149	mbk-1	W08G11.4	WBGene00012348	pptr-1
T04C10.1	WBGene00003149	mbk-1	Y37D8A.2	WBGene00012544	Y37D8A.2
T04C10.1	WBGene00003149	mbk-1	Y39A1A.21	WBGene00012658	Y39A1A.21
T04C10.1	WBGene00003149	mbk-1	Y39B6A.38	WBGene00012696	reps-1
T04C10.1	WBGene00003149	mbk-1	Y40B1A.4	WBGene00012735	sptf-3
T04C10.1	WBGene00003149	mbk-1	Y47D3A.28	WBGene00012935	mcm-10
T04C10.1	WBGene00003149	mbk-1	Y48G10A.2	WBGene00013019	Y48G10A.2
T04C10.1	WBGene00003149	mbk-1	Y51H1A.4	WBGene00013095	ing-3
T04C10.1	WBGene00003149	mbk-1	Y59A8B.22	WBGene00013354	snx-6
T04C10.1	WBGene00003149	mbk-1	Y66D12A.10	WBGene00013436	Y66D12A.10
T04C10.1	WBGene00003149	mbk-1	Y67H2A.10	WBGene00013465	Y67H2A.10
T04C10.1	WBGene00003149	mbk-1	ZK1098.4	WBGene00014221	ZK1098.4
T04C10.1	WBGene00003149	mbk-1	B0303.3	WBGene00015125	B0303.3
T04C10.1	WBGene00003149	mbk-1	BE0003N10.2	WBGene00015267	chin-1
T04C10.1	WBGene00003149	mbk-1	C16C8.11	WBGene00015849	C16C8.11
T04C10.1	WBGene00003149	mbk-1	C17E7.4	WBGene00015899	C17E7.4
T04C10.1	WBGene00003149	mbk-1	C25B8.4	WBGene00016088	clec-266
T04C10.1	WBGene00003149	mbk-1	C29E4.2	WBGene00016202	kle-2
T04C10.1	WBGene00003149	mbk-1	C30H7.2	WBGene00016278	C30H7.2
T04C10.1	WBGene00003149	mbk-1	C35E7.5	WBGene00016457	C35E7.5
T04C10.1	WBGene00003149	mbk-1	C35E7.6	WBGene00016458	C35E7.6
T04C10.1	WBGene00003149	mbk-1	C36C9.1	WBGene00016485	C36C9.1
T04C10.1	WBGene00003149	mbk-1	C41D11.7	WBGene00016566	eri-7
T04C10.1	WBGene00003149	mbk-1	F09F7.6	WBGene00017303	F09F7.6
T04C10.1	WBGene00003149	mbk-1	F17E9.4	WBGene00017541	F17E9.4
T04C10.1	WBGene00003149	mbk-1	F23H11.1	WBGene00017757	bra-2
T04C10.1	WBGene00003149	mbk-1	F26F4.5	WBGene00017827	F26F4.5
T04C10.1	WBGene00003149	mbk-1	F26F4.11	WBGene00017830	rpb-8
T04C10.1	WBGene00003149	mbk-1	F44E2.2	WBGene00018416	retr-1
T04C10.1	WBGene00003149	mbk-1	F45E12.1	WBGene00018474	cnep-1
T04C10.1	WBGene00003149	mbk-1	K09H11.3	WBGene00019600	rga-3
T04C10.1	WBGene00003149	mbk-1	R09A1.1	WBGene00019971	ergo-1
T04C10.1	WBGene00003149	mbk-1	R13F6.10	WBGene00020068	cra-1
T04C10.1	WBGene00003149	mbk-1	T08B2.5	WBGene00020346	T08B2.5
T04C10.1	WBGene00003149	mbk-1	T08B2.7	WBGene00020347	T08B2.7
T04C10.1	WBGene00003149	mbk-1	T22D1.10	WBGene00020687	ruvb-2
T04C10.1	WBGene00003149	mbk-1	W03A5.4	WBGene00020968	W03A5.4

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
T04C10.1	WBGene00003149	mbk-1	W03A5.6	WBGene00020970	W03A5.6
T04C10.1	WBGene00003149	mbk-1	W05F2.4	WBGene00021036	W05F2.4
T04C10.1	WBGene00003149	mbk-1	Y39G10AR.8	WBGene00021466	Y39G10AR.8
T04C10.1	WBGene00003149	mbk-1	Y44E3A.6	WBGene00021551	Y44E3A.6
T04C10.1	WBGene00003149	mbk-1	Y45G5AM.7	WBGene00021559	Y45G5AM.7
T04C10.1	WBGene00003149	mbk-1	Y50D4C.5	WBGene00021749	Y50D4C.5
T04C10.1	WBGene00003149	mbk-1	Y54F10BM.1	WBGene00021856	Y54F10BM.1
T04C10.1	WBGene00003149	mbk-1	Y55F3AM.1	WBGene00021920	mrps-25
T04C10.1	WBGene00003149	mbk-1	Y110A7A.8	WBGene00022458	prp-31
T04C10.1	WBGene00003149	mbk-1	ZK154.5	WBGene00022667	ZK154.5
T04C10.1	WBGene00003149	mbk-1	ZK1236.3	WBGene00023405	sor-1
T04C10.1	WBGene00003149	mbk-1	ZK662.4	WBGene00023497	lin-15B
T04C10.1	WBGene00003149	mbk-1	ZK678.1	WBGene00023498	lin-15A
T04C10.1	WBGene00003149	mbk-1	F11A10.8	WBGene00044329	cpsf-4
F49E11.1	WBGene00003150	mbk-2	K07C11.2	WBGene00000098	air-1
F49E11.1	WBGene00003150	mbk-2	C25A11.4	WBGene00000100	ajm-1
F49E11.1	WBGene00003150	mbk-2	C36A4.8	WBGene00000264	brc-1
F49E11.1	WBGene00003150	mbk-2	Y56A3A.20	WBGene00000369	ccf-1
F49E11.1	WBGene00003150	mbk-2	ZC518.3	WBGene00000376	ccr-4
F49E11.1	WBGene00003150	mbk-2	F01F1.8	WBGene00000381	cct-6
F49E11.1	WBGene00003150	mbk-2	F22D3.1	WBGene00000459	ceh-38
F49E11.1	WBGene00003150	mbk-2	Y54F10AM.4	WBGene00000464	ceh-44
F49E11.1	WBGene00003150	mbk-2	T25G3.2	WBGene00000496	chs-1
F49E11.1	WBGene00003150	mbk-2	R107.6	WBGene00000549	cls-2
F49E11.1	WBGene00003150	mbk-2	Y60A3A.10	WBGene00000987	dhs-24
F49E11.1	WBGene00003150	mbk-2	C18D1.1	WBGene00000995	die-1
F49E11.1	WBGene00003150	mbk-2	C25F6.2	WBGene00001006	dlg-1
F49E11.1	WBGene00003150	mbk-2	F43G6.1	WBGene00001016	dna-2
F49E11.1	WBGene00003150	mbk-2	T23G7.1	WBGene00001061	dpl-1
F49E11.1	WBGene00003150	mbk-2	Y102A5C.18	WBGene00001161	efl-1
F49E11.1	WBGene00003150	mbk-2	Y48C3A.17	WBGene00001162	efl-2
F49E11.1	WBGene00003150	mbk-2	F55A8.1	WBGene00001186	egl-18
F49E11.1	WBGene00003150	mbk-2	F52D2.4	WBGene00001569	gei-12
F49E11.1	WBGene00003150	mbk-2	M04B2.3	WBGene00001585	gfl-1
F49E11.1	WBGene00003150	mbk-2	ZC308.1	WBGene00001596	gld-2
F49E11.1	WBGene00003150	mbk-2	T07F8.3	WBGene00001597	gld-3
F49E11.1	WBGene00003150	mbk-2	Y18D10A.5	WBGene00001746	gsk-3
F49E11.1	WBGene00003150	mbk-2	ZK381.1	WBGene00001862	him-3
F49E11.1	WBGene00003150	mbk-2	ZK742.1	WBGene00002078	xpo-1
F49E11.1	WBGene00003150	mbk-2	F26F12.7	WBGene00002637	let-418
F49E11.1	WBGene00003150	mbk-2	F57B9.2	WBGene00002845	let-711
F49E11.1	WBGene00003150	mbk-2	ZK637.7	WBGene00002998	lin-9
F49E11.1	WBGene00003150	mbk-2	C03B8.4	WBGene00003002	lin-13
F49E11.1	WBGene00003150	mbk-2	K10B2.1	WBGene00003009	lin-23
F49E11.1	WBGene00003150	mbk-2	C32F10.2	WBGene00003020	lin-35
F49E11.1	WBGene00003150	mbk-2	T22A3.3	WBGene00003083	lst-1
F49E11.1	WBGene00003150	mbk-2	Y69A2AR.30	WBGene00003161	mdf-2
F49E11.1	WBGene00003150	mbk-2	ZC302.1	WBGene00003405	mre-11
F49E11.1	WBGene00003150	mbk-2	ZK1127.1	WBGene00003784	nos-2
F49E11.1	WBGene00003150	mbk-2	C18E3.7	WBGene00004093	ppw-1
F49E11.1	WBGene00003150	mbk-2	C18E9.3	WBGene00004105	szy-20
F49E11.1	WBGene00003150	mbk-2	W02D9.1	WBGene00004181	pri-2
F49E11.1	WBGene00003150	mbk-2	ZK675.1	WBGene00004208	ptc-1

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
F49E11.1	WBGene00003150	mbk-2	F31E3.3	WBGene00004340	rfc-4
F49E11.1	WBGene00003150	mbk-2	C50E10.4	WBGene00004945	sop-2
F49E11.1	WBGene00003150	mbk-2	F32A6.5	WBGene00006064	sto-2
F49E11.1	WBGene00003150	mbk-2	T05H10.5	WBGene00006734	ufd-2
F49E11.1	WBGene00003150	mbk-2	C06A5.7	WBGene00006823	unc-94
F49E11.1	WBGene00003150	mbk-2	F44F1.7	WBGene00006902	vet-6
F49E11.1	WBGene00003150	mbk-2	Y111B2A.22	WBGene00007027	ssl-1
F49E11.1	WBGene00003150	mbk-2	C13G3.3	WBGene00007554	pptr-2
F49E11.1	WBGene00003150	mbk-2	C14B1.5	WBGene00007576	dph-1
F49E11.1	WBGene00003150	mbk-2	C33G3.6	WBGene00007906	C33G3.6
F49E11.1	WBGene00003150	mbk-2	C39E9.12	WBGene00008035	C39E9.12
F49E11.1	WBGene00003150	mbk-2	C47E12.4	WBGene00008149	pyp-1
F49E11.1	WBGene00003150	mbk-2	D2005.5	WBGene00008400	drh-3
F49E11.1	WBGene00003150	mbk-2	F16A11.3	WBGene00008878	ppfr-1
F49E11.1	WBGene00003150	mbk-2	F17C11.10	WBGene00008921	F17C11.10
F49E11.1	WBGene00003150	mbk-2	F26H11.2	WBGene00009180	nurf-1
F49E11.1	WBGene00003150	mbk-2	F30A10.6	WBGene00009264	sac-1
F49E11.1	WBGene00003150	mbk-2	F36A2.13	WBGene00009460	F36A2.13
F49E11.1	WBGene00003150	mbk-2	F44F4.2	WBGene00009701	egg-3
F49E11.1	WBGene00003150	mbk-2	F54B3.1	WBGene00010013	F54B3.1
F49E11.1	WBGene00003150	mbk-2	F54B3.3	WBGene00010015	atad-3
F49E11.1	WBGene00003150	mbk-2	F54D5.5	WBGene00010051	F54D5.5
F49E11.1	WBGene00003150	mbk-2	H19N07.2	WBGene00010406	math-33
F49E11.1	WBGene00003150	mbk-2	H37N21.1	WBGene00010427	hpo-11
F49E11.1	WBGene00003150	mbk-2	K02B9.2	WBGene00010493	meg-2
F49E11.1	WBGene00003150	mbk-2	K04B12.2	WBGene00010550	K04B12.2
F49E11.1	WBGene00003150	mbk-2	K04C1.2	WBGene00010552	K04C1.2
F49E11.1	WBGene00003150	mbk-2	K09B11.9	WBGene00010713	uso-1
F49E11.1	WBGene00003150	mbk-2	M106.4	WBGene00010912	gmps-1
F49E11.1	WBGene00003150	mbk-2	R07E5.11	WBGene00011117	R07E5.11
F49E11.1	WBGene00003150	mbk-2	R07H5.8	WBGene00011128	R07H5.8
F49E11.1	WBGene00003150	mbk-2	T05F1.2	WBGene00011489	T05F1.2
F49E11.1	WBGene00003150	mbk-2	T05H10.1	WBGene00011507	T05H10.1
F49E11.1	WBGene00003150	mbk-2	T08A11.1	WBGene00011604	T08A11.1
F49E11.1	WBGene00003150	mbk-2	T12G3.1	WBGene00011737	sqst-1
F49E11.1	WBGene00003150	mbk-2	T24F1.2	WBGene00011994	samp-1
F49E11.1	WBGene00003150	mbk-2	T28C6.7	WBGene00012121	T28C6.7
F49E11.1	WBGene00003150	mbk-2	Y47D3A.28	WBGene00012935	mcm-10
F49E11.1	WBGene00003150	mbk-2	Y57G11C.3	WBGene00013301	Y57G11C.3
F49E11.1	WBGene00003150	mbk-2	Y105E8A.17	WBGene00013676	ekl-4
F49E11.1	WBGene00003150	mbk-2	Y106G6D.1	WBGene00013698	Y106G6D.1
F49E11.1	WBGene00003150	mbk-2	ZK512.5	WBGene00013985	sec-16
F49E11.1	WBGene00003150	mbk-2	ZK858.3	WBGene00014117	clec-91
F49E11.1	WBGene00003150	mbk-2	C02F5.4	WBGene00015347	cids-1
F49E11.1	WBGene00003150	mbk-2	C04E7.2	WBGene00015429	sor-3
F49E11.1	WBGene00003150	mbk-2	C05C8.9	WBGene00015466	hyls-1
F49E11.1	WBGene00003150	mbk-2	C06G3.6	WBGene00015552	C06G3.6
F49E11.1	WBGene00003150	mbk-2	C17E7.4	WBGene00015899	C17E7.4
F49E11.1	WBGene00003150	mbk-2	C17H12.9	WBGene00015934	ceh-48
F49E11.1	WBGene00003150	mbk-2	C30H7.2	WBGene00016278	C30H7.2
F49E11.1	WBGene00003150	mbk-2	C33H5.7	WBGene00016374	swd-2.2
F49E11.1	WBGene00003150	mbk-2	C35E7.6	WBGene00016458	C35E7.6
F49E11.1	WBGene00003150	mbk-2	C41H7.4	WBGene00016574	C41H7.4

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
F49E11.1	WBGene00003150	mbk-2	C49H3.4	WBGene00016791	C49H3.4
F49E11.1	WBGene00003150	mbk-2	C50F2.2	WBGene00016836	C50F2.2
F49E11.1	WBGene00003150	mbk-2	F09G2.8	WBGene00017316	F09G2.8
F49E11.1	WBGene00003150	mbk-2	F10E9.3	WBGene00017355	F10E9.3
F49E11.1	WBGene00003150	mbk-2	F13B9.1	WBGene00017419	F13B9.1
F49E11.1	WBGene00003150	mbk-2	F20D12.4	WBGene00017643	czw-1
F49E11.1	WBGene00003150	mbk-2	F35H10.7	WBGene00018072	nprl-3
F49E11.1	WBGene00003150	mbk-2	F39C12.1	WBGene00018193	F39C12.1
F49E11.1	WBGene00003150	mbk-2	F41H10.6	WBGene00018319	hda-6
F49E11.1	WBGene00003150	mbk-2	F44B9.8	WBGene00018409	F44B9.8
F49E11.1	WBGene00003150	mbk-2	F44E2.7	WBGene00018420	F44E2.7
F49E11.1	WBGene00003150	mbk-2	F49E8.1	WBGene00018635	nprl-2
F49E11.1	WBGene00003150	mbk-2	F55G1.9	WBGene00018904	F55G1.9
F49E11.1	WBGene00003150	mbk-2	H05C05.1	WBGene00019157	H05C05.1
F49E11.1	WBGene00003150	mbk-2	K07B1.5	WBGene00019465	acl-14
F49E11.1	WBGene00003150	mbk-2	K09H11.3	WBGene00019600	rga-3
F49E11.1	WBGene00003150	mbk-2	R09A1.1	WBGene00019971	ergo-1
F49E11.1	WBGene00003150	mbk-2	T10F2.2	WBGene00020422	T10F2.2
F49E11.1	WBGene00003150	mbk-2	T19B4.5	WBGene00020558	T19B4.5
F49E11.1	WBGene00003150	mbk-2	T21E3.1	WBGene00020652	egg-4
F49E11.1	WBGene00003150	mbk-2	T22D1.10	WBGene00020687	ruvb-2
F49E11.1	WBGene00003150	mbk-2	Y25C1A.6	WBGene00021293	Y25C1A.6
F49E11.1	WBGene00003150	mbk-2	Y32H12A.8	WBGene00021316	Y32H12A.8
F49E11.1	WBGene00003150	mbk-2	Y67D8C.5	WBGene00022069	eel-1
F49E11.1	WBGene00003150	mbk-2	Y77E11A.1	WBGene00022306	Y77E11A.1
F49E11.1	WBGene00003150	mbk-2	ZK1236.3	WBGene00023405	sor-1
F49E11.1	WBGene00003150	mbk-2	ZK662.4	WBGene00023497	lin-15B
F58B6.3	WBGene00003917	par-2	F48F7.1	WBGene00000105	alg-1
F58B6.3	WBGene00003917	par-2	Y56A3A.20	WBGene00000369	ccf-1
F58B6.3	WBGene00003917	par-2	T05G5.3	WBGene00000405	cdk-1
F58B6.3	WBGene00003917	par-2	F55H2.6	WBGene00000550	clu-1
F58B6.3	WBGene00003917	par-2	Y49A3A.5	WBGene00000877	cyn-1
F58B6.3	WBGene00003917	par-2	B0252.4	WBGene00000886	cyn-10
F58B6.3	WBGene00003917	par-2	K12H4.8	WBGene00000939	dcr-1
F58B6.3	WBGene00003917	par-2	K02F2.1	WBGene00001056	dpf-3
F58B6.3	WBGene00003917	par-2	R160.1	WBGene00001082	dpy-23
F58B6.3	WBGene00003917	par-2	M04B2.3	WBGene00001585	gfl-1
F58B6.3	WBGene00003917	par-2	ZC308.1	WBGene00001596	gld-2
F58B6.3	WBGene00003917	par-2	F22B7.13	WBGene00001688	gpr-1
F58B6.3	WBGene00003917	par-2	F08G12.10	WBGene00002124	inx-2
F58B6.3	WBGene00003917	par-2	Y43F4B.6	WBGene00002229	klp-19
F58B6.3	WBGene00003917	par-2	T19E10.1	WBGene00002297	ect-2
F58B6.3	WBGene00003917	par-2	B0261.2	WBGene00002583	let-363
F58B6.3	WBGene00003917	par-2	F28F8.3	WBGene00003079	lsm-5
F58B6.3	WBGene00003917	par-2	ZK858.4	WBGene00003209	mel-26
F58B6.3	WBGene00003917	par-2	W02A2.7	WBGene00003230	mex-5
F58B6.3	WBGene00003917	par-2	F45H11.2	WBGene00003587	ned-8
F58B6.3	WBGene00003917	par-2	K12D12.2	WBGene00003789	npp-3
F58B6.3	WBGene00003917	par-2	Y37E3.15	WBGene00003799	npp-13
F58B6.3	WBGene00003917	par-2	C15H11.3	WBGene00003834	nxf-1
F58B6.3	WBGene00003917	par-2	H39E23.1	WBGene00003916	par-1
F58B6.3	WBGene00003917	par-2	F54E7.3	WBGene00003918	par-3
F58B6.3	WBGene00003917	par-2	T26E3.3	WBGene00003921	par-6

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
F58B6.3	WBGene00003917	par-2	F09E5.1	WBGene00004034	pkc-3
F58B6.3	WBGene00003917	par-2	C18E9.3	WBGene00004105	szy-20
F58B6.3	WBGene00003917	par-2	K08D10.4	WBGene00004385	rnp-2
F58B6.3	WBGene00003917	par-2	F18E2.3	WBGene00004738	scc-3
F58B6.3	WBGene00003917	par-2	Y113G7A.3	WBGene00004754	sec-23
F58B6.3	WBGene00003917	par-2	Y116A8C.36	WBGene00006405	itsn-1
F58B6.3	WBGene00003917	par-2	R09B3.4	WBGene00006707	ubc-12
F58B6.3	WBGene00003917	par-2	Y40G12A.1	WBGene00006723	ubh-3
F58B6.3	WBGene00003917	par-2	K11E8.1	WBGene00006779	unc-43
F58B6.3	WBGene00003917	par-2	Y48B6A.3	WBGene00006964	xrn-2
F58B6.3	WBGene00003917	par-2	Y111B2A.22	WBGene00007027	ssl-1
F58B6.3	WBGene00003917	par-2	T04D1.4	WBGene00007053	chd-7
F58B6.3	WBGene00003917	par-2	C47G2.5	WBGene00008166	saps-1
F58B6.3	WBGene00003917	par-2	D2030.2	WBGene00008412	D2030.2
F58B6.3	WBGene00003917	par-2	D2045.2	WBGene00008422	D2045.2
F58B6.3	WBGene00003917	par-2	F10C2.4	WBGene00008645	F10C2.4
F58B6.3	WBGene00003917	par-2	F13D12.5	WBGene00008740	F13D12.5
F58B6.3	WBGene00003917	par-2	F43G9.1	WBGene00009664	idha-1
F58B6.3	WBGene00003917	par-2	F45E6.3	WBGene00009726	tbc-13
F58B6.3	WBGene00003917	par-2	F46B6.12	WBGene00009776	F46B6.12
F58B6.3	WBGene00003917	par-2	F58G1.1	WBGene00010263	wago-4
F58B6.3	WBGene00003917	par-2	K11E4.4	WBGene00010776	pix-1
F58B6.3	WBGene00003917	par-2	R07H5.2	WBGene00011122	cpt-2
F58B6.3	WBGene00003917	par-2	R53.2	WBGene00011272	R53.2
F58B6.3	WBGene00003917	par-2	T01G1.3	WBGene00011338	sec-31
F58B6.3	WBGene00003917	par-2	T05H10.1	WBGene00011507	T05H10.1
F58B6.3	WBGene00003917	par-2	T11F9.14	WBGene00011715	T11F9.14
F58B6.3	WBGene00003917	par-2	T12D8.2	WBGene00011730	drr-2
F58B6.3	WBGene00003917	par-2	T27A8.3	WBGene00012075	T27A8.3
F58B6.3	WBGene00003917	par-2	W04A4.5	WBGene00012234	W04A4.5
F58B6.3	WBGene00003917	par-2	Y49E10.2	WBGene00013029	glrx-5
F58B6.3	WBGene00003917	par-2	Y52D3.1	WBGene00013132	strd-1
F58B6.3	WBGene00003917	par-2	Y57A10A.26	WBGene00013266	Y57A10A.26
F58B6.3	WBGene00003917	par-2	Y113G7B.16	WBGene00013765	cdkr-3
F58B6.3	WBGene00003917	par-2	ZC412.3	WBGene00013884	ZC412.3
F58B6.3	WBGene00003917	par-2	ZK637.5	WBGene00014025	asna-1
F58B6.3	WBGene00003917	par-2	ZK1098.4	WBGene00014221	ZK1098.4
F58B6.3	WBGene00003917	par-2	ZK1320.9	WBGene00014258	ZK1320.9
F58B6.3	WBGene00003917	par-2	B0336.3	WBGene00015143	B0336.3
F58B6.3	WBGene00003917	par-2	B0403.4	WBGene00015168	pdi-6
F58B6.3	WBGene00003917	par-2	C01G5.6	WBGene00015308	C01G5.6
F58B6.3	WBGene00003917	par-2	C06E7.3	WBGene00015540	sams-4
F58B6.3	WBGene00003917	par-2	C28H8.9	WBGene00016200	dpff-1
F58B6.3	WBGene00003917	par-2	C34D4.14	WBGene00016405	hecd-1
F58B6.3	WBGene00003917	par-2	F01F1.4	WBGene00017161	rabn-5
F58B6.3	WBGene00003917	par-2	F07F6.4	WBGene00017217	F07F6.4
F58B6.3	WBGene00003917	par-2	F15B10.3	WBGene00017481	F15B10.3
F58B6.3	WBGene00003917	par-2	F40E3.2	WBGene00018230	mage-1
F58B6.3	WBGene00003917	par-2	F53H1.1	WBGene00018776	F53H1.1
F58B6.3	WBGene00003917	par-2	F53H1.4	WBGene00018778	F53H1.4
F58B6.3	WBGene00003917	par-2	F56C9.10	WBGene00018953	F56C9.10
F58B6.3	WBGene00003917	par-2	H24K24.3	WBGene00019240	H24K24.3
F58B6.3	WBGene00003917	par-2	R02D3.5	WBGene00019823	fnta-1

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
F58B6.3	WBGene00003917	par-2	R09A1.1	WBGene00019971	ergo-1
F58B6.3	WBGene00003917	par-2	R13F6.1	WBGene00020064	kbp-1
F58B6.3	WBGene00003917	par-2	Y34D9A.10	WBGene00021334	vps-4
F58B6.3	WBGene00003917	par-2	Y42H9AR.1	WBGene00021536	Y42H9AR.1
F58B6.3	WBGene00003917	par-2	Y48G8AL.5	WBGene00021686	Y48G8AL.5
F58B6.3	WBGene00003917	par-2	Y66H1A.2	WBGene00022044	dpm-1
F58B6.3	WBGene00003917	par-2	Y119D3B.11	WBGene00022488	orc-3
F58B6.3	WBGene00003917	par-2	ZK688.5	WBGene00022800	ZK688.5
T26E3.3	WBGene00003921	par-6	Y38C1AA.2	WBGene00000815	csn-3
T26E3.3	WBGene00003921	par-6	ZK520.4	WBGene00000837	cul-2
T26E3.3	WBGene00003921	par-6	F29D11.1	WBGene00003071	lrp-1
T26E3.3	WBGene00003921	par-6	K08A8.1	WBGene00003185	mek-1
T26E3.3	WBGene00003921	par-6	F54E7.3	WBGene00003918	par-3
T26E3.3	WBGene00003921	par-6	C10F3.5	WBGene00003954	pcm-1
T26E3.3	WBGene00003921	par-6	F09E5.1	WBGene00004034	pkc-3
T26E3.3	WBGene00003921	par-6	D2030.2	WBGene00008412	D2030.2
T26E3.3	WBGene00003921	par-6	F35B12.5	WBGene00009385	sas-5
T26E3.3	WBGene00003921	par-6	ZK1320.9	WBGene00014258	ZK1320.9
T26E3.3	WBGene00003921	par-6	C25B8.4	WBGene00016088	clec-266
T26E3.3	WBGene00003921	par-6	C30H7.2	WBGene00016278	C30H7.2
T26E3.3	WBGene00003921	par-6	F56F10.4	WBGene00018987	lgl-1
T26E3.3	WBGene00003921	par-6	Y38F2AR.5	WBGene00021423	tftc-1
C18G1.4	WBGene00003994	pgl-3	T05H4.13	WBGene00000110	alh-4
C18G1.4	WBGene00003994	pgl-3	C35D10.9	WBGene00000418	ced-4
C18G1.4	WBGene00003994	pgl-3	Y47G6A.8	WBGene00000794	crn-1
C18G1.4	WBGene00003994	pgl-3	R01H10.1	WBGene00001002	div-1
C18G1.4	WBGene00003994	pgl-3	K02F2.1	WBGene00001056	dpf-3
C18G1.4	WBGene00003994	pgl-3	M04B2.3	WBGene00001585	gfl-1
C18G1.4	WBGene00003994	pgl-3	T05A7.4	WBGene00001976	hmg-11
C18G1.4	WBGene00003994	pgl-3	F53A2.6	WBGene00002059	ife-1
C18G1.4	WBGene00003994	pgl-3	C05D11.12	WBGene00002855	let-721
C18G1.4	WBGene00003994	pgl-3	T27C10.6	WBGene00003068	lrk-1
C18G1.4	WBGene00003994	pgl-3	Y71G12B.20	WBGene00003111	mab-20
C18G1.4	WBGene00003994	pgl-3	T01G9.5	WBGene00003183	mei-1
C18G1.4	WBGene00003994	pgl-3	Y56A3A.1	WBGene00003826	ntl-3
C18G1.4	WBGene00003994	pgl-3	ZK381.4	WBGene00003992	pgl-1
C18G1.4	WBGene00003994	pgl-3	Y75B8A.30	WBGene00004085	pph-4.1
C18G1.4	WBGene00003994	pgl-3	Y110A7A.18	WBGene00004094	ppw-2
C18G1.4	WBGene00003994	pgl-3	D2030.6	WBGene00004178	prg-1
C18G1.4	WBGene00003994	pgl-3	K11E8.1	WBGene00006779	unc-43
C18G1.4	WBGene00003994	pgl-3	C04H5.1	WBGene00007312	C04H5.1
C18G1.4	WBGene00003994	pgl-3	C31C9.2	WBGene00007836	C31C9.2
C18G1.4	WBGene00003994	pgl-3	C36B1.3	WBGene00007971	rpb-3
C18G1.4	WBGene00003994	pgl-3	F14E5.2	WBGene00008800	F14E5.2
C18G1.4	WBGene00003994	pgl-3	F35B12.5	WBGene00009385	sas-5
C18G1.4	WBGene00003994	pgl-3	F58G1.1	WBGene00010263	wago-4
C18G1.4	WBGene00003994	pgl-3	M01E5.6	WBGene00010808	sepa-1
C18G1.4	WBGene00003994	pgl-3	R06C7.1	WBGene00011061	wago-1
C18G1.4	WBGene00003994	pgl-3	R07H5.2	WBGene00011122	cpt-2
C18G1.4	WBGene00003994	pgl-3	T09A5.9	WBGene00011637	sds-22
C18G1.4	WBGene00003994	pgl-3	T28C6.7	WBGene00012121	T28C6.7
C18G1.4	WBGene00003994	pgl-3	Y47D3B.9	WBGene00012943	bed-2
C18G1.4	WBGene00003994	pgl-3	ZK1098.1	WBGene00014218	ZK1098.1

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
C18G1.4	WBGene00003994	pgl-3	B0261.7	WBGene00015095	B0261.7
C18G1.4	WBGene00003994	pgl-3	C10G11.7	WBGene00015687	C10G11.7
C18G1.4	WBGene00003994	pgl-3	C34B2.6	WBGene00016391	C34B2.6
C18G1.4	WBGene00003994	pgl-3	C50E3.12	WBGene00016823	C50E3.12
C18G1.4	WBGene00003994	pgl-3	F53B1.4	WBGene00018737	F53B1.4
C18G1.4	WBGene00003994	pgl-3	H28G03.1	WBGene00019249	H28G03.1
C18G1.4	WBGene00003994	pgl-3	K10C2.4	WBGene00019620	fah-1
C18G1.4	WBGene00003994	pgl-3	R13F6.10	WBGene00020068	cra-1
C18G1.4	WBGene00003994	pgl-3	T09B4.9	WBGene00020383	T09B4.9
C18G1.4	WBGene00003994	pgl-3	W02D3.2	WBGene00020932	dhod-1
C18G1.4	WBGene00003994	pgl-3	Y4C6B.1	WBGene00021155	Y4C6B.1
C18G1.4	WBGene00003994	pgl-3	Y39G10AR.10	WBGene00021468	epg-2
C18G1.4	WBGene00003994	pgl-3	Y65B4BL.5	WBGene00022037	acs-13
C18G1.4	WBGene00003994	pgl-3	Y104H12D.2	WBGene00022426	Y104H12D.2
Y49E10.14	WBGene00004027	pie-1	T07D3.7	WBGene00000106	alg-2
Y49E10.14	WBGene00004027	pie-1	ZC308.1	WBGene00001596	gld-2
Y49E10.14	WBGene00004027	pie-1	T07F8.3	WBGene00001597	gld-3
Y49E10.14	WBGene00004027	pie-1	C53A5.3	WBGene00001834	hda-1
Y49E10.14	WBGene00004027	pie-1	W03C9.7	WBGene00003228	mex-1
Y49E10.14	WBGene00004027	pie-1	ZC302.1	WBGene00003405	mre-11
Y49E10.14	WBGene00004027	pie-1	C49H3.5	WBGene00003827	ntl-4
Y49E10.14	WBGene00004027	pie-1	C07G1.5	WBGene00004101	hgrs-1
Y49E10.14	WBGene00004027	pie-1	R09A8.3	WBGene00006493	tag-147
Y49E10.14	WBGene00004027	pie-1	W09C5.2	WBGene00006793	unc-59
Y49E10.14	WBGene00004027	pie-1	F35G2.2	WBGene00009436	marb-1
Y49E10.14	WBGene00004027	pie-1	F43G9.5	WBGene00009668	cfim-1
Y49E10.14	WBGene00004027	pie-1	T04D3.2	WBGene00011432	sdz-30
Y49E10.14	WBGene00004027	pie-1	Y17G7B.3	WBGene00012459	Y17G7B.3
Y49E10.14	WBGene00004027	pie-1	ZK1098.11	WBGene00014226	ZK1098.11
Y49E10.14	WBGene00004027	pie-1	B0041.2	WBGene00015007	ain-2
Y49E10.14	WBGene00004027	pie-1	C06G1.4	WBGene00015547	ain-1
Y49E10.14	WBGene00004027	pie-1	C30H7.2	WBGene00016278	C30H7.2
Y49E10.14	WBGene00004027	pie-1	C34B2.7	WBGene00016392	sdha-2
Y49E10.14	WBGene00004027	pie-1	C45E1.4	WBGene00016663	C45E1.4
Y49E10.14	WBGene00004027	pie-1	F41H10.6	WBGene00018319	hda-6
Y49E10.14	WBGene00004027	pie-1	F48E3.3	WBGene00018604	uggt-1
Y49E10.14	WBGene00004027	pie-1	W07E6.1	WBGene00021073	nol-1
M03D4.1	WBGene00006974	zen-4	H15N14.1	WBGene00000079	adr-1
M03D4.1	WBGene00006974	zen-4	F27C1.7	WBGene00000230	atp-3
M03D4.1	WBGene00006974	zen-4	T03E6.7	WBGene00000776	cpl-1
M03D4.1	WBGene00006974	zen-4	K08E3.6	WBGene00000875	cyk-4
M03D4.1	WBGene00006974	zen-4	F36H9.3	WBGene00000976	dhs-13
M03D4.1	WBGene00006974	zen-4	C25F6.2	WBGene00001006	dlg-1
M03D4.1	WBGene00006974	zen-4	F54D5.8	WBGene00001031	dnj-13
M03D4.1	WBGene00006974	zen-4	T05C3.5	WBGene00001037	dnj-19
M03D4.1	WBGene00006974	zen-4	T15H9.7	WBGene00001038	dnj-20
M03D4.1	WBGene00006974	zen-4	ZK328.2	WBGene00001166	eftu-2
M03D4.1	WBGene00006974	zen-4	C46A5.9	WBGene00001827	hcf-1
M03D4.1	WBGene00006974	zen-4	C53A5.3	WBGene00001834	hda-1
M03D4.1	WBGene00006974	zen-4	F26D10.3	WBGene00002005	hsp-1
M03D4.1	WBGene00006974	zen-4	C37H5.8	WBGene00002010	hsp-6
M03D4.1	WBGene00006974	zen-4	F53A2.6	WBGene00002059	ife-1
M03D4.1	WBGene00006974	zen-4	B0348.6	WBGene00002061	ife-3

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
M03D4.1	WBGene00006974	zen-4	F26B1.3	WBGene00002073	ima-2
M03D4.1	WBGene00006974	zen-4	C02F5.1	WBGene00002231	knl-1
M03D4.1	WBGene00006974	zen-4	Y71H2AM.19	WBGene00002244	laf-1
M03D4.1	WBGene00006974	zen-4	F26F12.7	WBGene00002637	let-418
M03D4.1	WBGene00006974	zen-4	T09A5.10	WBGene00002994	lin-5
M03D4.1	WBGene00006974	zen-4	T27C4.4	WBGene00003025	lin-40
M03D4.1	WBGene00006974	zen-4	C25D7.6	WBGene00003155	mcm-3
M03D4.1	WBGene00006974	zen-4	F26F4.7	WBGene00003598	nhl-2
M03D4.1	WBGene00006974	zen-4	T19B4.2	WBGene00003793	npp-7
M03D4.1	WBGene00006974	zen-4	C49H3.5	WBGene00003827	ntl-4
M03D4.1	WBGene00006974	zen-4	F48E8.5	WBGene00003901	paa-1
M03D4.1	WBGene00006974	zen-4	C17E4.5	WBGene00003904	pabp-2
M03D4.1	WBGene00006974	zen-4	C10F3.5	WBGene00003954	pcm-1
M03D4.1	WBGene00006974	zen-4	W03D2.4	WBGene00003955	pcn-1
M03D4.1	WBGene00006974	zen-4	Y37E3.9	WBGene00004014	phb-1
M03D4.1	WBGene00006974	zen-4	F45E4.2	WBGene00004046	plp-1
M03D4.1	WBGene00006974	zen-4	Y76A2B.1	WBGene00004075	pod-1
M03D4.1	WBGene00006974	zen-4	C07G1.5	WBGene00004101	hgrs-1
M03D4.1	WBGene00006974	zen-4	F44B9.7	WBGene00004125	mdt-30
M03D4.1	WBGene00006974	zen-4	T20H4.3	WBGene00004189	pars-1
M03D4.1	WBGene00006974	zen-4	F01G4.1	WBGene00004204	swsn-4
M03D4.1	WBGene00006974	zen-4	T23G5.1	WBGene00004391	rnr-1
M03D4.1	WBGene00006974	zen-4	F57B9.10	WBGene00004462	rpn-6.1
M03D4.1	WBGene00006974	zen-4	C52E4.4	WBGene00004501	rpt-1
M03D4.1	WBGene00006974	zen-4	Y49E10.1	WBGene00004506	rpt-6
M03D4.1	WBGene00006974	zen-4	W02B12.2	WBGene00004699	rsp-2
M03D4.1	WBGene00006974	zen-4	ZK652.1	WBGene00004918	snr-5
M03D4.1	WBGene00006974	zen-4	B0336.9	WBGene00006359	swp-1
M03D4.1	WBGene00006974	zen-4	F42A8.2	WBGene00006433	sdhb-1
M03D4.1	WBGene00006974	zen-4	T27E9.1	WBGene00006439	ant-1.1
M03D4.1	WBGene00006974	zen-4	C14B1.4	WBGene00006474	wdr-5.1
M03D4.1	WBGene00006974	zen-4	D1054.15	WBGene00006481	plrg-1
M03D4.1	WBGene00006974	zen-4	C36E8.5	WBGene00006537	tbb-2
M03D4.1	WBGene00006974	zen-4	F58A4.8	WBGene00006540	tbg-1
M03D4.1	WBGene00006974	zen-4	F39H11.2	WBGene00006577	tlf-1
M03D4.1	WBGene00006974	zen-4	Y92C3B.2	WBGene00006697	uaf-1
M03D4.1	WBGene00006974	zen-4	W02D3.9	WBGene00006773	unc-37
M03D4.1	WBGene00006974	zen-4	W09C5.2	WBGene00006793	unc-59
M03D4.1	WBGene00006974	zen-4	Y50E8A.4	WBGene00006795	unc-61
M03D4.1	WBGene00006974	zen-4	C27H6.2	WBGene00007784	ruvb-1
M03D4.1	WBGene00006974	zen-4	C36A4.5	WBGene00007966	C36A4.5
M03D4.1	WBGene00006974	zen-4	C37E2.1	WBGene00007993	idhb-1
M03D4.1	WBGene00006974	zen-4	F11A10.2	WBGene00008683	repo-1
M03D4.1	WBGene00006974	zen-4	F25D7.4	WBGene00009113	F25D7.4
M03D4.1	WBGene00006974	zen-4	F32B6.2	WBGene00009319	mccc-1
M03D4.1	WBGene00006974	zen-4	F35G2.2	WBGene00009436	marb-1
M03D4.1	WBGene00006974	zen-4	F38E11.5	WBGene00009542	copb-2
M03D4.1	WBGene00006974	zen-4	F44E5.1	WBGene00009688	F44E5.1
M03D4.1	WBGene00006974	zen-4	F45G2.4	WBGene00009732	cope-1
M03D4.1	WBGene00006974	zen-4	F53A2.7	WBGene00009952	F53A2.7
M03D4.1	WBGene00006974	zen-4	F54B3.3	WBGene00010015	atad-3
M03D4.1	WBGene00006974	zen-4	K08E3.5	WBGene00010665	K08E3.5
M03D4.1	WBGene00006974	zen-4	M18.5	WBGene00010890	ddb-1

Bait_ID	Bait_WB_ID	Bait_symbol	Prey_ID	Prey_WB_ID	Prey_symbol
M03D4.1	WBGene00006974	zen-4	R07E5.3	WBGene00011111	snfc-5
M03D4.1	WBGene00006974	zen-4	T10B10.2	WBGene00011679	ucr-2.2
M03D4.1	WBGene00006974	zen-4	T20G5.1	WBGene00011867	chc-1
M03D4.1	WBGene00006974	zen-4	Y54G9A.6	WBGene00013209	bub-3
M03D4.1	WBGene00006974	zen-4	Y54G11A.13	WBGene00013220	ctl-3
M03D4.1	WBGene00006974	zen-4	Y57G11C.34	WBGene00013324	mrps-7
M03D4.1	WBGene00006974	zen-4	ZK669.4	WBGene00014054	ZK669.4
M03D4.1	WBGene00006974	zen-4	B0303.3	WBGene00015125	B0303.3
M03D4.1	WBGene00006974	zen-4	B0403.4	WBGene00015168	pdi-6
M03D4.1	WBGene00006974	zen-4	C25B8.4	WBGene00016088	clec-266
M03D4.1	WBGene00006974	zen-4	C44B7.10	WBGene00016630	C44B7.10
M03D4.1	WBGene00006974	zen-4	F20D12.4	WBGene00017643	czw-1
M03D4.1	WBGene00006974	zen-4	F38A5.7	WBGene00018164	sup-36
M03D4.1	WBGene00006974	zen-4	F44B9.8	WBGene00018409	F44B9.8
M03D4.1	WBGene00006974	zen-4	F48E3.3	WBGene00018604	uggt-1
M03D4.1	WBGene00006974	zen-4	F49D11.1	WBGene00018625	prp-17
M03D4.1	WBGene00006974	zen-4	F56F11.4	WBGene00018991	F56F11.4
M03D4.1	WBGene00006974	zen-4	F57B10.14	WBGene00019007	F57B10.14
M03D4.1	WBGene00006974	zen-4	H28G03.1	WBGene00019249	H28G03.1
M03D4.1	WBGene00006974	zen-4	R07G3.5	WBGene00019941	pgam-5
M03D4.1	WBGene00006974	zen-4	R119.3	WBGene00020089	R119.3
M03D4.1	WBGene00006974	zen-4	T10F2.4	WBGene00020423	prp-19
M03D4.1	WBGene00006974	zen-4	T19B4.5	WBGene00020558	T19B4.5
M03D4.1	WBGene00006974	zen-4	T20F5.6	WBGene00020628	T20F5.6
M03D4.1	WBGene00006974	zen-4	T22D1.10	WBGene00020687	ruvb-2
M03D4.1	WBGene00006974	zen-4	T24C4.1	WBGene00020757	ucr-2.3
M03D4.1	WBGene00006974	zen-4	Y39G10AR.2	WBGene00021460	zwl-1
M03D4.1	WBGene00006974	zen-4	Y47G6A.24	WBGene00021648	mis-12
M03D4.1	WBGene00006974	zen-4	Y55F3AR.3	WBGene00021934	cct-8
M03D4.1	WBGene00006974	zen-4	Y61A9LA.8	WBGene00022019	sut-2
M03D4.1	WBGene00006974	zen-4	Y65B4A.6	WBGene00022029	Y65B4A.6
M03D4.1	WBGene00006974	zen-4	Y71F9AL.9	WBGene00022114	Y71F9AL.9
M03D4.1	WBGene00006974	zen-4	ZK616.4	WBGene00022774	swsn-6

VII.4. Publications

Book chapter

- Sury MD*, **Chen JX***, Selbach M. (2014) In Vivo Stable Isotope Labeling by Amino Acids in Drosophila melanogaster. *Methods in Molecular Biology* 1188:85-93. (*shared first author)

Manuscript under review

- **Chen JX**, Cipriani PG, Mecenas D, Polanowska J, Piano F, Gunsalus KC, Selbach M (2015) *In vivo* protein interaction proteomics in *C. elegans* embryos provides new insights into P granule dynamics.

Published articles during PhD thesis period

- Wang Y, Wang J, Devaraj A, Singh M, Orgaz AJ, **Chen JX**, Selbach M, Ivics Z, Izsvák Z (2014) Suicidal Autointegration of Sleeping Beauty and piggyBac Transposons in Eukaryotic Cells. *PLoS Genetics* 10:e1004103.
- Sury MD*, **Chen JX***, Selbach M (2010) The SILAC fly allows for accurate protein quantification in vivo. *Molecular & Cellular Proteomics* 9:2173-83. (*shared first author)

Published articles prior to PhD thesis period

- Xu H, **Chen JX**, Zhang TM, Gong J, Wu QL, Wang JP (2007) Correlation between hand and total body bone density in normal Chinese children. *Bone* 41:360-5.
- Xu H, **Chen JX**, Gong J, Zhang TM, Wu QL, Yuan ZM, Wang JP (2007) Normal reference for bone density in healthy Chinese children. *Journal of Clinical Densitometry* 10:266-75.
- Xu H, Gong J, **Chen JX**, Zhang TM, Wu QL (2007) Bilateral femoral bone mineral density measurements in Chinese women and men. *Journal of Clinical Densitometry* 10:165-9.
- Zhang TM, Xu H, Yuan ZM, **Chen JX**, Gong J, Wu QL (2007) Assessment of total body fat percentage from regional spine and femur DXA measurements among Chinese women and men. *Journal of Clinical Densitometry* 10:55-64.
- **Chen JX**, Gong J, Zhang TM, Wu QL, Xu H (2006) Correlations between bone mineral density of the hand and other skeletal sites as measured by DXA in Chinese women and men. *Journal of Clinical Densitometry* 9:461-8.
- **Chen JX**, Zhang TM, Lim FL, Wu HC, Lei TF, Yeong PK, Xia SJ (2006) Current knowledge and attitudes about organ donation and transplantation among Chinese university students. *Transplantation Proceedings* 38:2761-5.

Berlin, den 22.09.2015

Jiaxuan Chen

VII.5. Meeting abstracts

- *In vivo* interaction mapping in *C. elegans* embryos. **Proteomic Forum 2015**. Berlin, Germany. (poster)
- *In vivo* interaction proteomics identifies GEI-12 as an MBK-2 interaction partner required for P-granule formation in *C. elegans* embryos. **SignGene Symposium 2014**. Berlin, Germany. (poster)
- *In vivo* interaction proteomics identifies GEI-12 as an MBK-2 interaction partner required for P-granule formation in *C. elegans* embryos. **Berlin Summer Meeting 2014**. Berlin, Germany. (poster)
- Quantitative interaction proteomics identifies GEI-12 as an MBK-2 interaction partner required for P-granule formation in *C. elegans*. **Berlin *C. elegans* Meeting 2014**. Berlin, Germany. (poster)
- Analysis of protein-protein interaction by *in vivo* quantitative proteomics during *C. elegans* embryogenesis. **EuPA 2013 Scientific Meeting**. Saint Malo, France. (poster)
- Analysis of protein-protein interaction by *in vivo* quantitative proteomics during *C. elegans* embryogenesis. **19th International *C. elegans* Meeting**. Los Angeles, USA. (poster)
- Analysis of Protein-Protein Interaction by *in vivo* Quantitative Proteomics during *C. elegans* Embryogenesis. **Proteomic Forum 2013**. Berlin, Germany. (poster)
- Analysis of Protein-Protein Interaction by *in vivo* Quantitative Proteomics in *C. elegans*. **13th MDC/FMP PhD Retreat 2011**. Liebenwalde, Germany. (poster)

VIII Acknowledgements

First and foremost, I would like to express my deepest gratitude to Prof. Matthias Selbach for guiding me through these years of wonderful scientific adventure. Matthias' passion for new knowledge, rigorous scientific attitude and insightful views on scientific questions inspired me to explore the unknowns and develop my research skills for becoming a scientist. It was a tremendous pleasure working closely with you for the last many years.

I had the honor to have two more advisors Prof. Fabio Piano and Prof. Kris Gunsalus at New York University who led me into the colorful and fascinating world of worm biology, warmly welcomed me as member of their labs and helped me with important advice on numerous occasions. My PhD project would not have been possible without the initiative from Fabio and Prof. Nikolaus Rajewsky on global PhD education and the generous support from the MDC-NYU PhD program. I enjoyed every moment working in NYU and I am proud of having being part of it. I would also like to thank Prof. Richard Lucius for supervising my doctoral thesis at Humboldt-Universität zu Berlin.

In the MDC, lab work would not have been so smooth without the wonderful assistance from Christian and Martha. Being a "fossil" member of the Selbach group, I have the privilege to have worked with my beloved colleagues: Björn, Flo, Fabian, Olivia, This, Marie, Erik, Djordje, Katrin, Henrik, Koshi, Katrina, Boris, Kamila, Michal, João and Murphy, as well as master's students and trainees: Sarbani, Anna-Laura, Murat, Piotr, Bastian, David, Karen, Burcu and Jonas. I thank you all for the entertaining scientific interactions, countless times of help on and off work, and friendship. I enjoyed the fun time playing with all the lovely future scientists during a recent baby boom: Liam, Martha, Kento, Toma, Ida, Elias and Friedemann. Also, I have to mention our "silent" heroes who work 24×7 for all my precious samples: Ernie, Bert, Tiffy, Kermit, Bibo, Grobi, Statler, Waldorf and Gladys.

In New York, I am very thankful to have had the great technical support from Jessica, Katherine, Sherly, Noah and Olivia. Special thanks go to Miyeko who kindly shared with me many of her beautiful GFP worms, Desirea for producing the GEI-12::GFP strain and Giselle for some of the RNAi assays. I would also like to thank other members of the Gunsiano lab: Sean, Kahn, Jola, Jerome, Indrani, Helene and Huey-Ling. It was a great pleasure working with all of you.

I am grateful to Jennifer Stewart who has helped me dealing with administrative and logistic work since day one of my PhD. I am much thankful to Sylvia Sibilak for taking the pressure of facing bureaucracy off my shoulder. I would also like to acknowledge Petra, Sabine, Carrie, Vinny and Jessica for their administrative assistance on both continents.

I met Matthias for the very first time in his lectures and the trademark saliva proteome lab training (the day when his Nature paper got accepted). Starting for my master's thesis, I've worked in Matthias' lab over

these years. It's been a very long journey full of fun and memorable moments: from the lab office entertained with constant input of music to my luxurious one-man satellite office, from being the first non-German speaker to now one of our small United Nations, from the calming green nature of Buch to the flamboyant high-rises of Manhattan, the drama of hurricane Sandy, the many occasions of looking into the microscope and thinking I have found something new. The list goes on and on. There were ups and downs of course, but I enjoyed my PhD years immensely alongside colleagues with the simple goal of doing good science.

Finally, I am indebted to my parents and sister for their endless love and support throughout all these years.

IX Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Berlin, den 22.09.2015

Jiaxuan Chen